

תרגיל מסכם – DATA ANALYST

Contents

1	תרגיל מסכם – DATA ANALYST	
1	תקציר מנהלים – על הנתונים	
2	חלק 1 – SQL	7
2.1	יש לפתח שלילפה שתציג את כמות דוחות החנייה לפי רובעים (Borough) בעיריית ניו יורק בין השנים 2015 – 2017	10
2.2	יש להוסיף לשלילפה הקודמת את היום בשבוע שבו ניתנו דוחות החנייה כך שהשלילפה תציג את כמות דוחות החנייה לכל רובע ולכל יום בשבוע.	12
2.3	יש לפתח שלילפה שתציג את חמשת סוגי עבירות החנייה, על-פי קוד עבירה ViolationCode, הכי נפוצות בעיריית ניו יורק. עד 2015 בשנים	13
2.4	יש להציג את שתי סוגי העבירות הכי נפוצות לכל צבע רכב בעיריית ניו יורק בשנים 2015 - 2017	15
2.5	יש לבנות שלילפה המציגה כמות הרכבים שקיבלו דוחות חניה בין השנים 2015 ל-2017 לפי קבוצות של:	16
2.6	יש להציג שלילפה המציגה לכל מדינה שבה רשום הרכב את העמודות הבאות:	17
3	חלק 2 – BI	19
3.1	דוח 1	19
3.1.1	מה הם 5 סוגי עבירות החנייה הנפוצות ביותר לאורך השנים? האם חל שינוי משנה לשנה	19
3.1.2	באיזה יום בשבוע ישנן יותר עבירות חניה? האם יש הבדל בין הרובעים השונים?	19
3.1.3	באילו שעות של היום, בפרקי זמן של שעות, יש יותר עבירות חניה? האם זה תלוי ברובע?	20
3.1.4	הציגו ויזואליזציה המראה כמה רכבים ביצעו יותר מ-10 עבירות חניה, כמה בין 5-9 וכמה מתחת ל-5 עבירות	22
3.2	דוח 2	22
3.2.1	מהו סוג הרכב, המקבל הכי הרבה דוחות חנייה ב-NY?	22
3.2.2	מהו גובה הקנס הממוצע לכל סוג רכב או לקבוצת סוגי רכב	23
3.2	האם יש צבע רכב דומיננטי?	24
3.2.3	מאיזו מדינה State, באים רוב הרכבים המעורבים בעבירות חניה?	24
3.3	דוח 3	25
3.3.1	הציגו על מפה גיאוגרפית את עבירות החניה שבוצעו עם יכולת סינון של שעות היום והלילה. (הציגו חלוקה הגיונית של השעות ליום ולילה) (השתמשו בשדה Address עבור זיהוי המקומות על גבי המפה) (האם אפשר לזהות בצורה ברורה אזורים ו/או חלקי יום בעייתיים)	25
3.3.2	מה הן עשרת הרחובות במנהטן שיש בהם הכי הרבה דוחות חניה	29
3.3.3	מה גובה הקנס הממוצע שמקבלים בעלי רכב אלו? (לא רק העשרה העליונים אלא כלל המשאיות)	30
3.4	דוח 4	31
3.4.1	מה ההכנסה, הרווח ואחוז הרווח (רווח חלקי הכנסות) של עיריית NY מאכיפת חוקי החנייה בכל שנה לכל סוכנות	32
3.4.2	מה הרובע borough, הכי רווחי בעיר NYC?	33
3.4.3	מה הם חמשת סוגי הקנס הרווחיים ביותר	34
4	המלצות:	35
5	סיכום אישי	38
6	תיאור המערכת	39

1. תקציר מנהלים – על הנתונים

כשמסתכלים על הנתונים, ניתן בצורה מובהקת לראות ש NY זאת עיר עבודה, בייחוד מנהטן. רוב הדוחות שמתקבלים בעיר זה בזמן העבודה, בעוד שביללות הדוחות יורדים בכתשעים אחוזים

	TOTAL	סטטן	ברוקלין	קוינס	מנהטן	ברונקס	
לילה	87698	727	K23	K19	K31	14K	
יום	686818	7947	K173	K146	K289	K146	
	12.8%	9.2%	13.3%	13%	10.7%	9.6%	

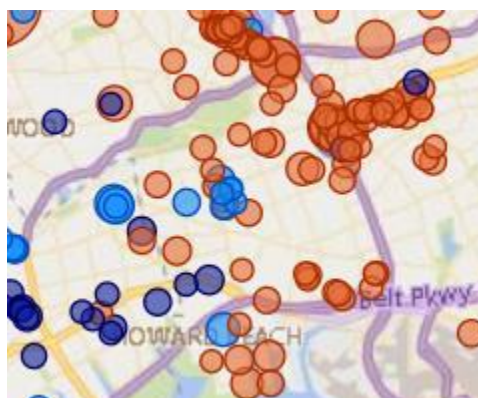
יותר מכך : ניכר שבסופי שבוע וחגים כשאין עבודה – גם כמות הדוחות יורדים בצורה משמעותית ניתן לראות שרוב מוחלט מהדוחות ניתן לתושבי העיר , אך גם לספקי עבודה שמתנייעים לעיר בשעות הכבודה ולמשאיות ואוטובוסים

רואים שהעיר מחולקת ל 5 רובעים, (מנהטן ברונקס קווינס ברוקלין וסטטן איילנד)

על פי איכות הנתונים, נראה שהם נרשמים למערכת בצורה ידנית, ויש מקום לשיפור בנושא. טעויות בהכנסת נתונים הן בעוכריה של כל מערכת, וההשפעה היא בעיקר על למידה מהמידע על מנת לשפר תהליכים יש מקומות של אנומליות וחוסר בהירות : למשל רואים שאין כמעט רכבים של העירייה שמקבלים דוחות, האם זה אומר שהם לא מבצעים עבירות ? או שאינם מקבלים עבירות רק בגלל שהם חברים של מישהו ? פקח אולי ?

Government 3

כמו כן רואים שהרבה סוגי מכוניות לא נקלטים נכון, ודוחות מדווחים ברבעים לא נכונים למשל : מה פשר הערבובים הללו בין ברונקס ומנהטן ? יש נקודות שממש לא שייכות שמפוזרות במפה



נדרשנו לזהות את סוגי המכונות לפי סוג המרכב, הבעיה שהנתון הזה נזיל מאוד – הוא תלוי בזיהוי של הפקח, במקרה הטוב (לא תמיד יש את הנתון בכלל) הפקח ישתמש במערכת זיהוי (Ticket Issuing Device (TID) ואם הוא מעדיף להכניס לבד , זה מסתכם בדו משמעויות ... ויש הרבה מקומות שהנתון יכול להיות שגוי או רב משמעי ...

אני שמת'י לב שיש נתון נוסף plate type : שהוא יותר משמעותי לדעתי : 1. הוא לא חסר
2. הוא קשור במספר של הרכב 3. המספר שלו (הרשיון לעסוק במשהו) מספר על המימוש שלו, ולא כמו במצב השני, שהצורה שלו מספרת על מי שהוא , אולי

ברגע שמקישים את מספר הרכב – המערכת אמורה לזהות גם את המרכב של הרכב.
רכב יכול להיות עם מרכב מסוים – אבל זה לא בהכרח מגדיר את השימוש שלו, מה שבהחלט מאופיין ב plate type

אי אפשר להסתמך על שניהם במערכת - כי כשניסיתי להוסיף את זה למודל, ולחבר עם fact table זה התנגד לי ולא יכולתי להמשיך ... כנראה בגלל שלפעמים לא הוקשו ערכים נכונים במערכת

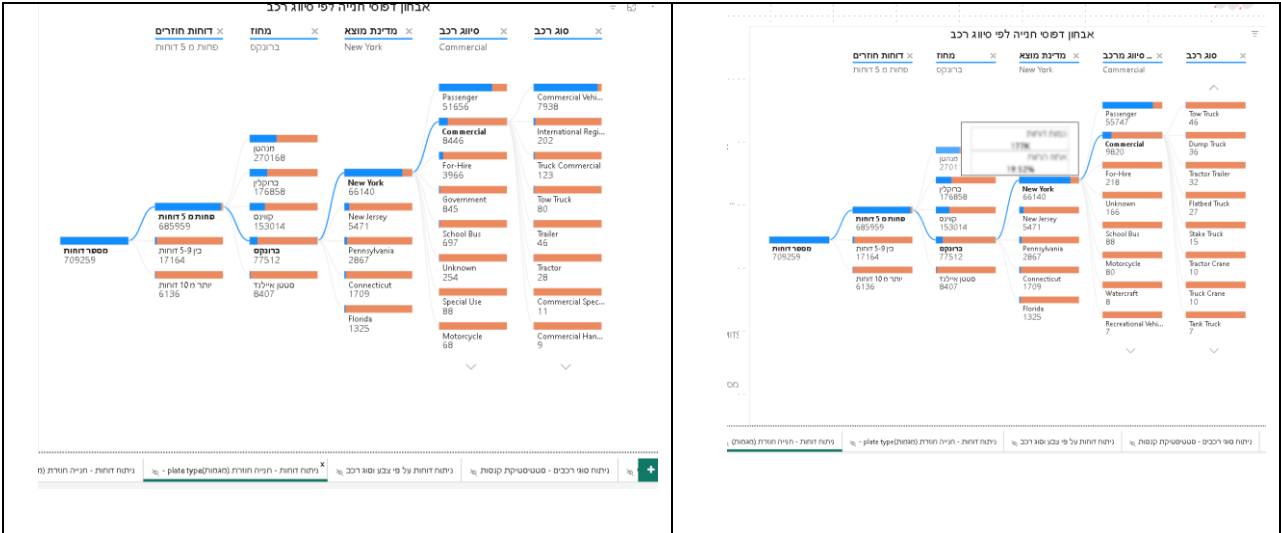
בבסיס הנתונים שקיבלנו היו חסרים מאות קודים שנמצאו במערכת (של body type) לעומת ה plate type שהיו כולם בפנים

ביצעתי סיווג (classification) בשני סוגי הנתונים ולשניהם היו אותם ערכי סיווג בסופו של דבר, מה שמלמד על קורלציה

בסופו של דבר עבדתי עם ה body type כי זה מה שדרש הפרויקט, אבל הייתי מעדיף לעבוד עם plate type בפעם הבאה

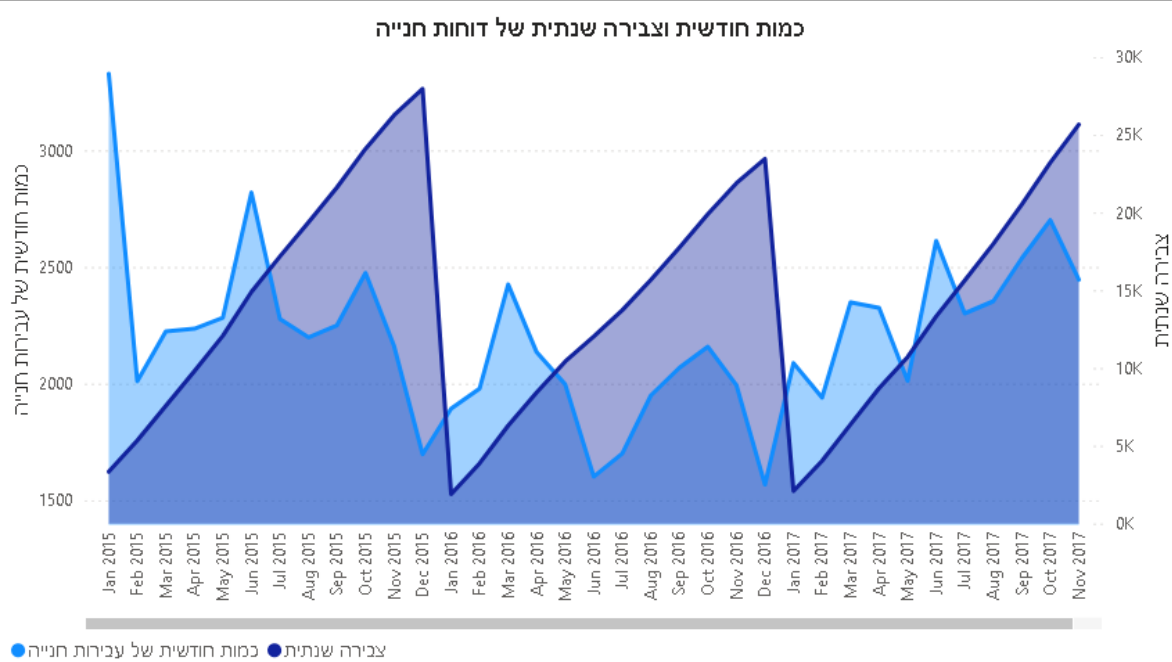
לדוגמא אותו מסך , בפעם ראשונה משתמש ב plate type ובפעם השנייה ב body type (שלטענתי פחות מדויק , גם כי הוא לא שלם וגם בגלל שאפשר להכניס אותו ידנית והוא לא תמיד משקף את השימוש האמיתי)

body type	לעומת שימוש ב plate type – :
-----------	------------------------------



במערכת היו 943 ערכים של body type שמתוכם 192 לא מזוהים, ולאף אחד אין סיווג אמיתי התחלתי להכין את טבלת הסיווג תוך שימוש במערכת הסיווג של משרד התחבורה, אך משום שהזמן דחק – ביצעתי כמה פקודות מובנות ב chat gpt על מנת ללמוד ולהבין את כולם וניסיתי למנוע הזיות ככל האפשר (שאלתה ב fact table חשפה אותם, ונשאר לי לסווג את כולם וגם לחשוף מה המשמעות של כל אלו שלא במערכת ...)

דבר מעניין קרה לי, כשצפיתי בנתונים החודשיים (בגרף הבא) :
זה גרף שמראה את הצבירה החודשית, כנגד הצבירה השנתית שמחושבת כל שנה מחדש (מצטברת שנתית)

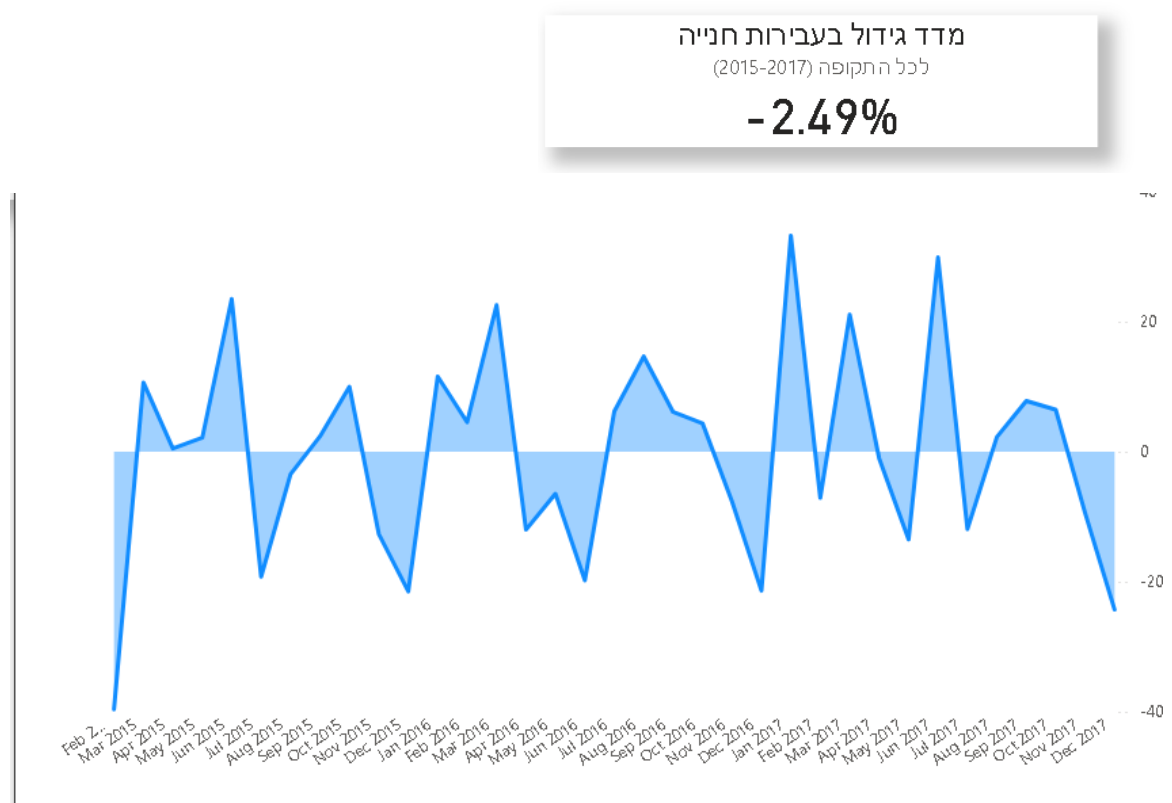


אנחנו רואים שבכל שנה (בתכלת) יש ירידות משמעותיות בנובמבר – דצמבר
את הפיקים אפשר להסביר בסוף רבעונים ואולי לפני סוף שנת מס – לפני היציאה לחופש
תמיד לפני ירידה ממש תלולה תהייה עלייה מאוד חדה, אין מתינות
מה שמעניין זה שבפברואר 2015 יש ממש ירידה חדה, וזה מוזר כי רבעון נגמר במרץ ובפברואר בשנים
אחרות רואים מגמות עלייה דווקא

מסתבר שבפברואר 2015 היה גל קור שפקד את ניו יורק מינוס 33 מעלות
כנראה אנים הגיעו לעבודה (או לא, אני לא יודע), אבל פקחים לא היו

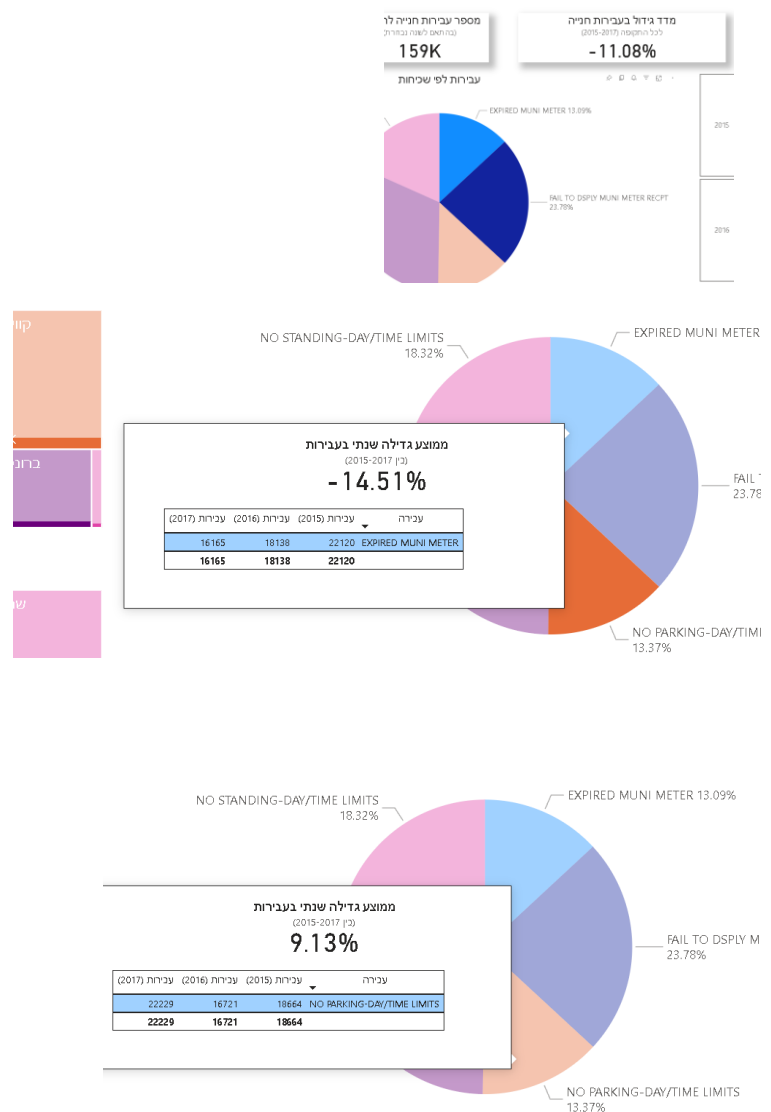
דבר נוסף שאינו מובהק:

יש ירידה בכמות הדוחות בין 2015 ל 2017 (לא משמעותית ואולי אפילו בתוך סטיית תקן, בייחוד
בהסתמך על זה שהנתונים הוכנסו ידנית, אז זה לא מובהק)



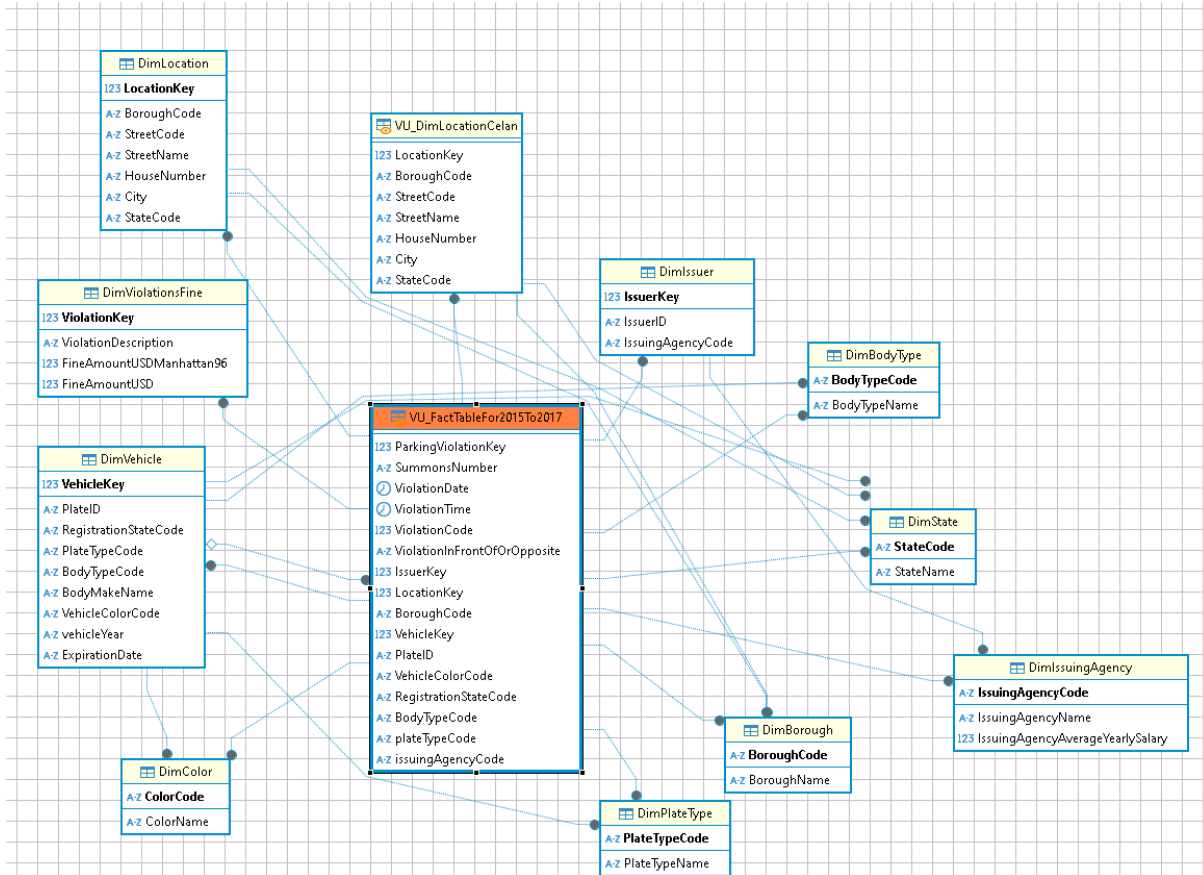
בעצם ניתן לראות בגרף המצורף את ההפרשים בין החודשים בין הדוחות
אם נשווה את השטח שמתחת ל 0 לשטח שמעליו, כנראה נראה את הירידה

מה שכן : ניתן לראות שיש ירידות חדות בחלק מהסוגים של הדוחות, בייחוד כאלו של חוסר תשומת לב (בתמונה רואים את העבירות של מדחן שהסתיים ושל מדבקה שלא הומגה כראוי) בירידה של 11% ביחד, בזמן שעברות שקשורות לחנייה במקומות שאסור לעמוד בהם עלו, יכל להיות שניתן לראות את זה לשייך זאת לסוגי הרכבים שעשו אותם, דבר שלא בדקתי, מפאת הזמן, אבל אני מנחש שזאת עבירה שקשורה לעבודה (פריקה/איסוף של סחורה ונוסעים) .. בהחלט דבר שאבדוק במשך



רוב האנשים שעוברים עבירות הם מתוך NY, לא אנשים מבחוץ
 ויש קבוצה מצומצמת שעושה המון יחסית לגודלה : commercial

2. חלק 1 – SQL



הערות מקדימות:

- כשנסתכלים על ה DDL רואים ש ParkingViolationKey הוא PK אחרת חייבים להשתמש ב distinct בשליפת fpv.ParkingViolationKey

- יש במערכת 4296 boroughs שהם unknown החלטתי לא להתייחס אליהם, משום שהם משמעותיים בסטטיסטיקה החניות, והמצאותם רק מזיקה לתמונה הכללית, ולא היו הנחיות בקשר אליהם יצרתי view ללא הנתון הזה ואני פונה אליו בכל השאלות הרלוונטיות אם ההנחיה תהייה כן להתייחס לזה, אשנה את השאלתה ב VIEW ולא אצטרך לשנות בכל הפניות לטבלה:

VU_DimLocationClean

```
CREATE VIEW dbo.VU_DimLocationClean AS
SELECT  LocationKey
        ,BoroughCode
        ,StreetCode
        ,StreetName
        ,HouseNumber
        ,City
        ,StateCode
FROM DWH_DATA_ANALYST.dbo.DimLocation
WHERE BoroughCode <> 99 --99 = unknown borough
;
```

Data Analyst 09/2024
עידו ביסטרי 025704149
נייד: 0523578697

- כיתה'ך מיטוב הנתונים, אני שמתי לב שבכל הפעמים רוצים שליפה עבור שנים 2015-2017 (חוץ משאלה 2, אבל אני מניח שזה חסר בגלל טעות), אז מראש הכנתי VIEW ללא הנתונים הללו וגם החלפתי שם את הסוג של השדה IssueDate ל DATE
- יצרתי שדה אחד שמראה את ה dates ואחד את הזמן - השדות ViolationTime, ViolationDate (לא חיברתי אותם לשדה אחד כי לדעתי זה יותר נוח למידול ודורש פחות משאבים מ POWER BI)
- הנחתי שאנו רוצים לראות רכבים ללא צבע, משום שזה לא פרמטר שמפריע לסטטיסטיקות המרחביות (יש K217 כאלו בבסיס הנתונים, אם לא נרצה לראות אותם, אז שינוי קטן בקוד של ה view יאפשר עבודה איתם – יש להוסיף את השורה בסוף הקוד :
(and dv.VehicleColorCode <> 'UNK';)
- על מנת לאפשר מבנה של star model וכדי למנוע גלישה ל now flake , הוספתי (בעצת מיישה) את הkeys הרלוונטיים מטבלאות המימדים (dim tables) לתוך ה VIWE של טבלת העובדות (fact table).
- באותה צורה הנחתי שאנו לא רוצים לראות עבירות ללא מחוז כי זה מפריע לסטטיסטיקה (ראה הערה קודמת לגבי מיימד המחוזות)

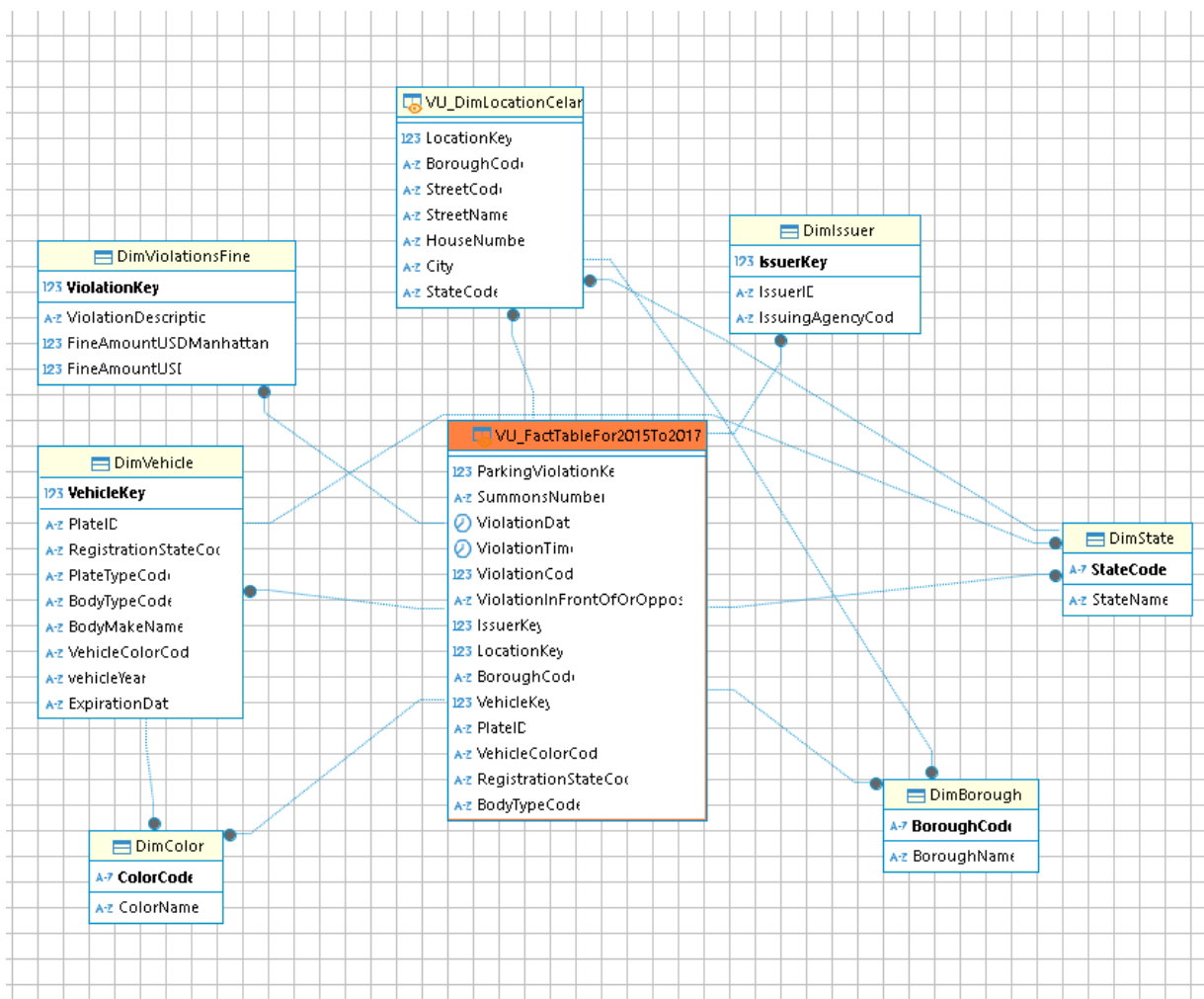
VU_FactTableFor2015To2017

```
CREATE VIEW dbo.VU_FactTableFor2015To2017
AS
SELECT ParkingViolationKey
, SummonsNumber
, CAST(fpv.IssueDate AS DATE) AS ViolationDate
, TRY_CAST(
    LEFT(fpv.ViolationTime, 2) + ':' + SUBSTRING(ViolationTime, 3, 2) +
    ' ' +
    CASE
        WHEN RIGHT(fpv.ViolationTime, 1) = 'A' THEN 'AM'
        WHEN RIGHT(fpv.ViolationTime, 1) = 'P' THEN 'PM'
    END
    AS TIME) AS ViolationTime
-- ,CAST(IssueDate AS DATETIME) +
-- COALESCE(
--     TRY_CAST(
--     STUFF(LEFT(ViolationTime, 4), 3, 0, ':') + -- Insert ":" at position 3
--     ' ' +
--     CASE
--         WHEN RIGHT(ViolationTime, 1) = 'A' THEN 'AM'
--         WHEN RIGHT(ViolationTime, 1) = 'P' THEN 'PM'
--     END
--     AS datetime),
--     '1900-01-01 00:00:00' -- Default invalid values to 00:00 AM
-- ) AS ViolationDateTime
, fpv.ViolationCode
, fpv.ViolationInFrontOfOrOpposite
, fpv.IssuerKey
, fpv.LocationKey
, vdlc.BoroughCode
, dv.VehicleKey
, dv.PlateID
, dv.VehicleColorCode
, dv.RegistrationStateCode
, dv.BodyTypeCode
, dv.plateTypeCode
, di.issuingAgencyCode
FROM FactParkingViolation fpv
    join VU_DimLocationClean vdlc on vdlc.LocationKey = fpv.LocationKey
    join DimVehicle dv on dv.VehicleKey = fpv.VehicleKey
    left join DimIssuer di on di.IssuerKey = fpv.IssuerKey
WHERE YEAR(CAST(fpv.IssueDate AS DATE)) BETWEEN 2015 AND 2017
--and dv.VehicleColorCode <> 'UNK';
;
```

בכל השליפות וגם ב BI אשתמש ב VIEW במקום בטבלה המקורית

2.1 יש לפתח שליפה שתציג את כמות דוחות החנייה לפי רובעים (Borough) בעיריית ניו יורק בין השנים 2015 – 2017

- יש להציג את שם הרובע.
- יש למיין את התוצאות לפי סדר יורד של כמות דוחות החנייה
- בסיום פיתוח השאילתה יש להפוך אותה ל- Stored Procedure כך שהפרוצדורה תופעל עם פרמטר של שם הרובע.



SQL:

```
SELECT BoroughName
, count (vfpv.ParkingViolationKey ) AS #tickets
FROM DimBorough db
JOIN VU_DimLocationClean vdlc
ON vdlc.BoroughCode = db.BoroughCode
JOIN VU_FactTableFor2015To2017 vfpv
ON vdlc.LocationKey = vfpv.LocationKey
GROUP BY BoroughName
ORDER BY 2 desc;
```

Results 1		
	A-Z BoroughName	123 #tickets
1	Manhattan	319,750
2	Brooklyn	195,616
3	Queens	165,630
4	Bronx	84,846
5	Staten Island	8,674

STORED PROC:

```
CREATE PROCEDURE GetBoroughTicketCounts @Borough NVARCHAR(255) = NULL AS BEGIN
    SET NOCOUNT ON;

    SELECT BoroughName
           ,count (vfpv.ParkingViolationKey ) AS #tickets
    FROM DimBorough db
    JOIN VU_DimLocationClean vdlc
        ON vdlc.BoroughCode = db.BoroughCode
    JOIN VU_FactTableFor2015To2017 vfpv
        ON vdlc.LocationKey = vfpv.LocationKey
    WHERE (@Borough IS NULL OR LOWER(db.BoroughName) LIKE LOWER(@Borough)+'%')
    GROUP BY BoroughName
    ORDER BY #tickets DESC;
END;
```

EXEC [dbo].[GetBoroughTicketCounts] @borough='b'

Grid		Az BoroughName	123 #tickets
1		Brooklyn	195,616
2		Bronx	84,846

2.2 יש להוסיף לשליפה הקודמת את היום בשבוע שבו ניתנו דוחות החנייה כך שהשליפה תציג את כמות דוחות החנייה לכל רובע ולכל יום בשבוע.

- יש להציג את שם היום בשבוע (לא את המספר)
- תוצאת השליפה תהיה ממוינת לפי רובע ויום בשבוע
- בסיום פיתוח השאילתה יש להפוך אותה ל- Stored Procedure כך שהפרוצדורה תופעל עם פרמטר של יום בשבוע

הערות:

אין כאן דרישה להגבלה בשנים 2015-2017, אבל רשום להוסיף לשאילתה הקודמת שכן מסננת את זה

SQL:

```
SELECT BoroughName,
       DATENAME(WEEKDAY, vftft.ViolationDateTime) AS DayName,
       --DATEPART(WEEKDAY, vftft.ViolationDateTime) AS _daynum,
       count(vftft.ParkingViolationKey) AS #tickets
FROM DimBorough db
     JOIN VU_DimLocationClean vdlc
       ON vdlc.BoroughCode = db.BoroughCode
     JOIN VU_FactTableFor2015To2017 vftft
       ON vftft.LocationKey = vdlc.LocationKey
GROUP BY BoroughName,
         DATENAME(WEEKDAY, vftft.ViolationDateTime) ,
         DATEPART(WEEKDAY, vftft.ViolationDateTime)
ORDER BY BoroughName,
         DATEPART(WEEKDAY, vftft.ViolationDateTime) -- number of
```

Results 1 X				
SELECT BoroughName, DATENAME(WEEKDAY, vftft.IssueDate), DATEPART(WEEKDAY, vftft.IssueDate) AS _daynum, count(vftft.ParkingViolationKey) AS #tickets				
	A-Z BoroughName	A-Z DayName	123 _daynum	123 #tickets
1	Bronx	Sunday	1	4,173
2	Bronx	Monday	2	13,203
3	Bronx	Tuesday	3	14,813
4	Bronx	Wednesday	4	12,951
5	Bronx	Thursday	5	15,017
6	Bronx	Friday	6	14,926
7	Bronx	Saturday	7	9,763
8	Brooklyn	Sunday	1	7,477
9	Brooklyn	Monday	2	30,194
10	Brooklyn	Tuesday	3	36,233
11	Brooklyn	Wednesday	4	31,044
12	Brooklyn	Thursday	5	36,197
13	Brooklyn	Friday	6	33,212
14	Brooklyn	Saturday	7	21,259
15	Manhattan	Sunday	1	11,896
16	Manhattan	Monday	2	48,764

- מספיק האותיות הראשונות ברובע
- מספיק האותיות הראשונות של היום

StoredProc:

```
CREATE PROCEDURE GetParkingViolationCountsForBoroughAndDayOfWeek
    @Day NVARCHAR(50) = NULL, -- Day name parameter (e.g., 'Monday')
    @Borough NVARCHAR(255) = NULL -- Borough name parameter (e.g., 'Manhattan')
AS
select BoroughName, DayName , #tickets
from (
    SELECT BoroughName,
           DATENAME(WEEKDAY, vftft.ViolationDateTime) AS DayName,
           DATEPART(WEEKDAY, vftft.ViolationDateTime) AS _daynum,
           count(vftft.ParkingViolationKey) AS #tickets
    FROM DimBorough db
         JOIN VU_DimLocationClean vdlc
           ON vdlc.BoroughCode = db.BoroughCode
         JOIN VU_FactTableFor2015To2017 vftft
           ON vftft.LocationKey = vdlc.LocationKey
    GROUP BY BoroughName,
             DATENAME(WEEKDAY, vftft.ViolationDateTime) ,
             DATEPART(WEEKDAY, vftft.ViolationDateTime)
) sql4DayAndBoroughName
where (@Borough is null or sql4DayAndBoroughName.BoroughName like lower(@Borough)+'%')
and (@Day is null or sql4DayAndBoroughName.DayName like lower(@Day)+'%')

ORDER BY BoroughName,
         _daynum-- number of day
```

EXEC [dbo].[GetParkingViolationCountsForBoroughAndDayOfWeek]

Results 1

EXEC [dbo].[GetParkingViolationCountsForBoroughAndDayOfWeek] Enter a SQL expression to filter

	A-Z BoroughName	A-Z DayName	123 #tickets
1	Bronx	Sunday	4,173
2	Bronx	Monday	13,433
3	Bronx	Tuesday	14,913
4	Bronx	Wednesday	12,951
5	Bronx	Thursday	15,017
6	Bronx	Friday	14,926
7	Bronx	Saturday	9,763
8	Brooklyn	Sunday	7,477
9	Brooklyn	Monday	30,194
10	Brooklyn	Tuesday	36,233
11	Brooklyn	Wednesday	31,044
12	Brooklyn	Thursday	36,197
13	Brooklyn	Friday	33,212
14	Brooklyn	Saturday	21,259
15	Manhattan	Sunday	11,586
16	Manhattan	Monday	48,764
17	Manhattan	Tuesday	52,212

EXEC [dbo].[GetParkingViolationCountsForBoroughAndDayOfWeek] @borough='staten island'

Results 1

EXEC [dbo].[GetParkingViolationCountsForBoroughAndDayOfWeek] Enter a SQL expression to filter results (use Ctrl+Space)

	A-Z BoroughName	A-Z DayName	123 #tickets
1	Staten Island	Sunday	562
2	Staten Island	Monday	1,336
3	Staten Island	Tuesday	1,392
4	Staten Island	Wednesday	1,473
5	Staten Island	Thursday	1,426
6	Staten Island	Friday	1,436
7	Staten Island	Saturday	1,049

EXEC [dbo].[GetParkingViolationCountsForBoroughAndDayOfWeek] @Day='sunday'

Results 1

EXEC [dbo].[GetParkingViolationCountsForBoroughAndDayOfWeek] Enter a SQL expression to filter results (use Ctrl+Sp

	A-Z BoroughName	A-Z DayName	123 #tickets
1	Bronx	Sunday	4,173
2	Brooklyn	Sunday	7,477
3	Manhattan	Sunday	11,586
4	Queens	Sunday	5,594
5	Staten Island	Sunday	562

EXEC [dbo].[GetParkingViolationCountsForBoroughAndDayOfWeek] @Day='sunday', @borough='q'

Results 1

EXEC [dbo].[GetParkingViolationCountsForBoroughAndDayOfWeek] Enter a SQL expression to filter

	A-Z BoroughName	A-Z DayName	123 #tickets
1	Queens	Sunday	5,594

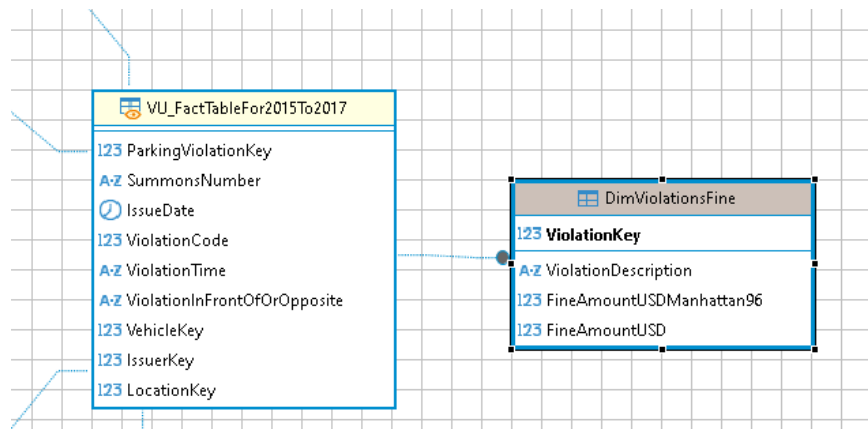
2.3 יש לפתח שליפה שתציג את חמשת סוגי עבירות החנייה, על-פי קוד עבירה ViolationCode, הכי נפוצות בעיריית ניו יורק. עד 2017 ו-2015 בשנים

- בסיום פיתוח השאלתה יש להפוך אותה ל- Stored Procedure
- כשהפרוצדורה תופעל עם פרמטר של מספר העבירות הכי נפוצות

יבאתי (imported) את טבלת הקנסות מאקסל לבסיס הנתונים על מנת ליצור אחידות בסביבת העבודה מאוחר יותר אשתמש בטבלה גם ב POWER BI, אלא אם כן הדרישה תהייה אחרת

בוחר את הנתונים בטבלה אני רואה כפילויות של נתונים – אחד לאחד (distinct זה בדיוק חצי) אז אני מעיף את כל הכפולים על ידי יצירת טבלה חדש, ואז מייצר PK עבור ה column וגם משנה את ה type מ varchar(50) ל int גם ב FACT TABLE

```
use DWH_DATA_ANALYST
SELECT DISTINCT * INTO DimViolationsFine FROM DimFine;
ALTER TABLE FactParkingViolation ALTER COLUMN VIOLATIONCODE int ;
ALTER TABLE DimViolationsFine ALTER COLUMN VIOLATION_CODE int NOT NULL;
ALTER TABLE DimViolationsFine ADD CONSTRAINT PK_DimViolations_ViolationCode PRIMARY KEY CLUSTERED (VIOLATION_CODE);
-- rename the name of the columns
EXEC DWH_DATA_ANALYST.sys.sp_rename N'DWH_DATA_ANALYST.dbo.DimViolationsFine.VIOLATION_CODE', N'ViolationKey', 'COLUMN';
drop table dimFine
```



SQL:

```
SELECT top(5) COUNT(vftft.ParkingViolationKey) AS count#
,dvf.ViolationDescription
FROM VU_FactTableFor2015To2017 vftft
JOIN DimViolationsFine dvf
ON dvf.ViolationKey = vftft.ViolationCode
GROUP BY dvf.ViolationDescription
ORDER BY count# desc
```

Results 1		
SELECT top(5) COUNT(vftft.ParkingViolationKey) AS count#, dvf.ViolationDescription		
Grid	123 count#	A-Z ViolationDescription
1	135,945	NO PARKING-STREET CLEANING
2	102,489	FAIL TO DPLY MUNI METER RECPT
3	79,190	NO STANDING-DAY/TIME LIMITS
4	57,819	NO PARKING-DAY/TIME LIMITS
5	56,427	EXPIRED MUNI METER

STORED PROC:

```
CREATE PROCEDURE GetTopParkingViolations
@TopNumber INT = 0 -- Default is 0, meaning no Limit ,
-- IF WE EXECUTE IT WITHOUT PARAMETER THE SAME AS 0
AS
BEGIN
SET NOCOUNT ON;
DECLARE @SQL NVARCHAR(MAX);

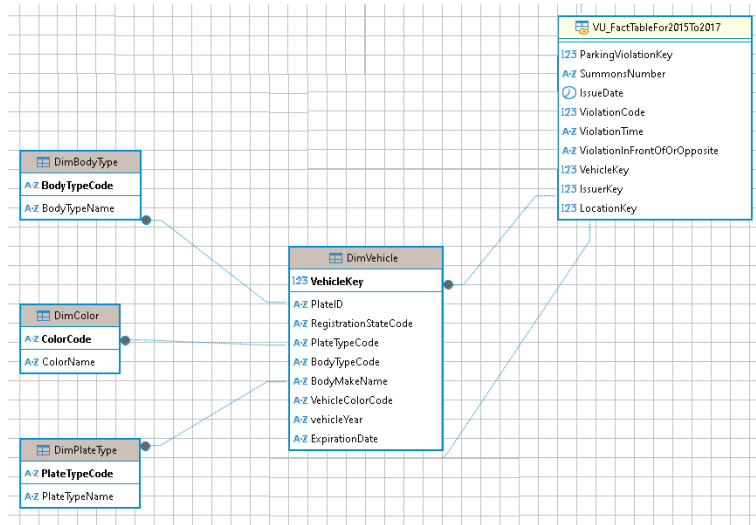
-- create DYNAMIC SQL
SET @SQL =
'SELECT ' +
CASE WHEN @TopNumber > 0
THEN
'TOP (' + CAST(@TopNumber AS NVARCHAR) + ') '
ELSE
''
END +
'COUNT(vftft.ParkingViolationKey) AS count#
,dvf.ViolationDescription
FROM VU_FactTableFor2015To2017 vftft
JOIN DimViolationsFine dvf
ON dvf.ViolationKey = vftft.ViolationCode
GROUP BY dvf.ViolationDescription
ORDER BY count# desc';

-- Execute the DYNAMIC SQL
EXEC sp_executesql @SQL;
END;
```

EXEC [dbo].[GetTopParkingViolations] @TopNumber=8		
Results 1		
EXEC [dbo].[GetTopParkingViolations] @TopNumber=8		
Grid	123 count#	A-Z ViolationDescription
1	135,945	NO PARKING-STREET CLEANING
2	102,489	FAIL TO DPLY MUNI METER RECPT
3	79,190	NO STANDING-DAY/TIME LIMITS
4	57,819	NO PARKING-DAY/TIME LIMITS
5	56,427	EXPIRED MUNI METER

2.4 יש להציג את שתי סוגי העבירות הכי נפוצות לכל צבע רכב בעיריית ניו יורק בשנים 2015 - 2017

- יש להימנע מלהציג צבע רכב לא ידוע.
- בסיום פיתוח השאילתה יש להפוך אותה ל- Stored Procedure כשהפרוצדורה תופעל עם פרמטר של מספר העבירות הכי נפוצות



SQL:

```

SELECT top(5)
COUNT(vftft.ParkingViolationKey) AS count#
,dc.ColorName
FROM VU_FactTableFor2015To2017 vftft
JOIN DimVehicle dv
    ON vftft.VehicleKey = dv.VehicleKey
JOIN DimColor dc
    ON dc.ColorCode = dv.VehicleColorCode
WHERE dv.VehicleColorCode <> 'UNK'
GROUP BY dv.VehicleColorCode
    
```

Results 1		
	count#	ColorName
1	235,643	WHITE
2	62,898	BLUE
3	62,250	BLACK
4	56,347	GREY
5	44,555	RED

STORED PROC:

```
-- STORED PROCEDURE
CREATE PROCEDURE GetTopParkingViolationsByColor
    @TopNumber INT = 0 -- Default is 0, meaning no Limit ,
                        -- IF WE EXECUTE IT WITHOUT PARAMETER THE SAME AS 0
AS
BEGIN
    SET NOCOUNT ON;
    DECLARE @SQL NVARCHAR(MAX);

    -- create DYNAMIC SQL
    SET @SQL =
        'SELECT ' +
        CASE WHEN @TopNumber > 0
        THEN
            'TOP (' + CAST(@TopNumber AS NVARCHAR) + ') '
        ELSE
            ''
        END +
        'COUNT(vftft.ParkingViolationKey) AS count#
        ,dc.ColorName
        FROM VU_FactTableFor2015To2017 vftft
        JOIN DimVehicle dv
            ON vftft.VehicleKey = dv.VehicleKey
        JOIN DimColor dc
            ON dc.ColorCode = dv.VehicleColorCode
        WHERE dv.VehicleColorCode <> ''UNK''
        GROUP BY dv.VehicleColorCode
        ,dc.ColorName
        ORDER BY count# DESC';

    -- Execute the DYNAMIC SQL
    EXEC sp_executesql @SQL;
END;
```

2.5 יש לבנות שליפה המציגה כמות הרכבים שקיבלו דוחות חניה בין השנים 2015 ל 2017 לפי קבוצות של:

- 10 או יותר
- 5-9
- פחות מ 5

SQL:

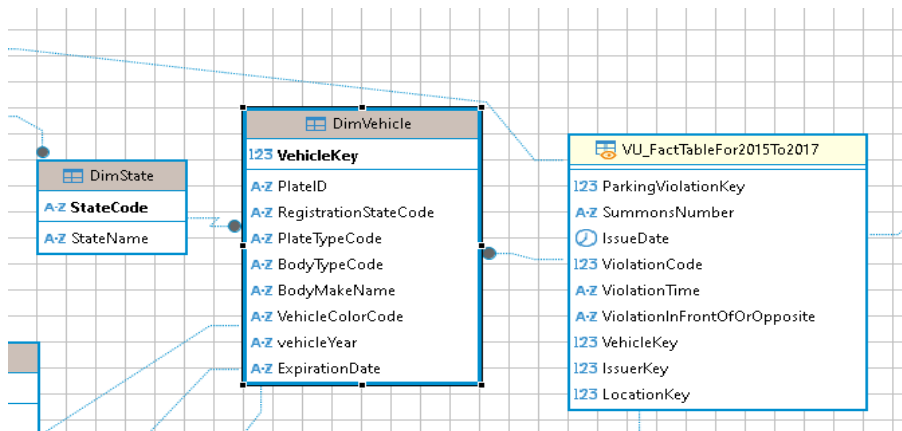
```
with parking_violations_cte (VehicleKey,count#,ViolationsAmountClassification)
AS
    (select vftft.VehicleKey ,
        count(vftft.ParkingViolationKey) as count# ,
        case
            when count(vftft.ParkingViolationKey) > 9 then 'group of 10 or more violations'
            when count(vftft.ParkingViolationKey) > 4 then 'group of 5-9 violations'
            else 'group with less then 5 violations'
        end as ViolationClassification
        from VU_FactTableFor2015To2017 vftft
        group by vftft.VehicleKey
    )
select count(vehiclekey), ViolationsAmountClassification
from parking_violations_cte
group by ViolationsAmountClassification
```

100 %		
Results Messages		
	(No column name)	ViolationsAmountClassification
1	4589	group of 10 or more violations
2	8156	group of 5-9 violations
3	532089	group with less then 5 violations

2.6 יש להציג שליפה המציגה לכל מדינה שבה רשום הרכב את העמודות הבאות:

- כמות דוחות החנייה בשנת 2015
- כמות דוחות החנייה בשנת 2016
- כמות דוחות החנייה בשנת 2017

•אחוז השינוי של כמות דוחות החנייה בין שנת 2017 לבין שנת 2015 (יש להציג את המספר באחוזים)



SQL:

```
WITH
year_data_cte
AS
(
    SELECT
        COUNT(vftft.ParkingViolationKey) AS tickets,
        ds.StateName AS state,
        YEAR(vftft.ViolationDateTime) AS Year
    FROM
        VU_FactTableFor2015To2017 vftft
    JOIN
        DimVehicle dv ON vftft.VehicleKey = dv.VehicleKey
    JOIN
        DimState ds ON dv.RegistrationStateCode = ds.StateCode
    GROUP BY
        ds.StateName, YEAR(vftft.ViolationDateTime)
)
SELECT
    state,
    value2015,
    value2016,
    value2017,
    CASE
        WHEN value2015 IS NOT NULL AND value2015 != 0 THEN
            ((value2017 - value2015) * 100.0 / value2015)
        ELSE
            NULL
    END AS PercentageChange
FROM (
    SELECT
        state,
        tickets,
        -- Get values for each year
        MAX(CASE WHEN year = 2015 THEN tickets END) OVER (PARTITION BY state) AS value2015,
        MAX(CASE WHEN year = 2016 THEN tickets END) OVER (PARTITION BY state) AS value2016,
        MAX(CASE WHEN year = 2017 THEN tickets END) OVER (PARTITION BY state) AS value2017
    FROM year_data_cte
) AS ranked_data
GROUP BY state, value2015, value2016, value2017
ORDER BY state
```

Results 1 x					
WITH year_data_cte AS (SELECT COUNT(vftft.ParkingViolationKey) AS tickets, ds.StateName AS state, YEAR(vftft.ViolationDateTime) AS Year FROM VU_FactTableFor2015To2017 vftft JOIN DimVehicle dv ON vftft.VehicleKey = dv.VehicleKey JOIN DimState ds ON dv.RegistrationStateCode = ds.StateCode GROUP BY ds.StateName, YEAR(vftft.ViolationDateTime)) AS ranked_data GROUP BY state, value2015, value2016, value2017 ORDER BY state					
	state	value2015	value2016	value2017	PercentageChange
1	Alabama	192	201	191	-0.5208333333
2	Alaska	29	34	17	-41.3793103448
3	Arizona	738	610	689	-6.6395663957
4	Arkansas	78	88	57	-26.9230769231
5	California	605	587	677	11.9008264463
6	Colorado	123	103	107	-13.0081300813
7	Connecticut	3,882	3,487	3,586	-7.6249356002
8	Delaware	390	371	385	-1.2820512821

3. חלק 2 – BI

3.1 דוח 1

3.1.1 מה הם 5 סוגי עבירות החנייה הנפוצות ביותר לאורך השנים? האם חל שינוי משנה לשנה

	Violation Type	2015	2016	2017	2015-2017
1	No Parking - Street Cleaning	30.50%	31.33%	32.54%	31.44%
2	Fail to Display Muni Meter Receipt	25.24%	23.81%	22.19%	23.78%
3	No Standing - Day/Time Limits	18%	18.05%	18.90%	18.32%
4	No Parking - Day/Time Limits	12.02%	12.86%	15.27%	13.37%
5	Expired Muni Meter	14.24%	13.95%	11.10%	13.09%

ניתן לראות שהעבירה הרביעית והחמישית התחלפו במיקום בין 2015 – 2016

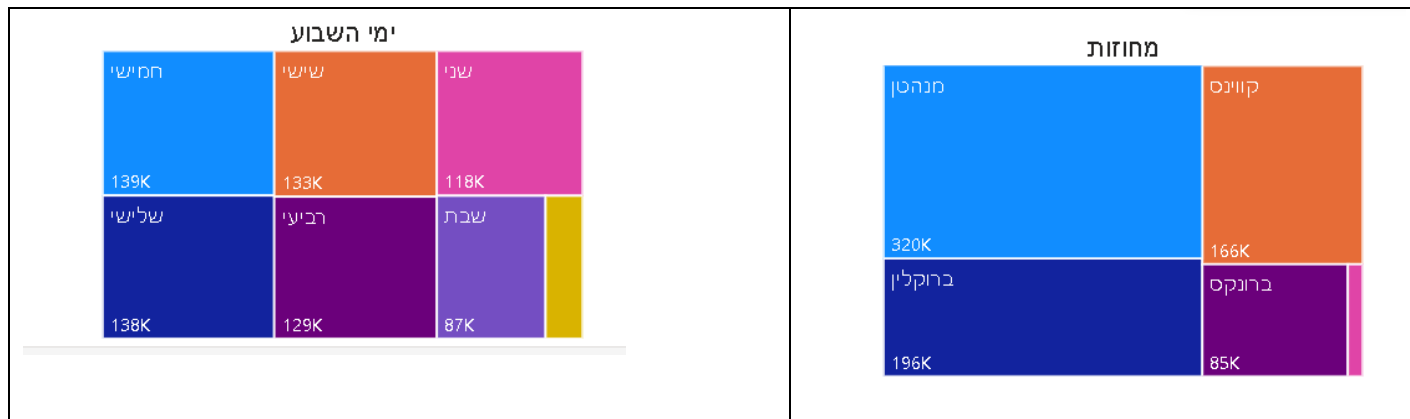
ניתן לראות שהעבירה שהכי חוזרת על עצמה היא שחנו בזמן ניקוי הכביש
 זה שירות שיש לחשוב עליו – הקשר על האזרח חשוב, ויכול להיות שעל בסיס תקשורת אנשים
 חונים באזורים אסורים זמנית (צריך לבדוק את זה)

3.1.2 באיזה יום בשבוע ישנן יותר עבירות חנייה? האם יש הבדל בין הרובעים השונים?

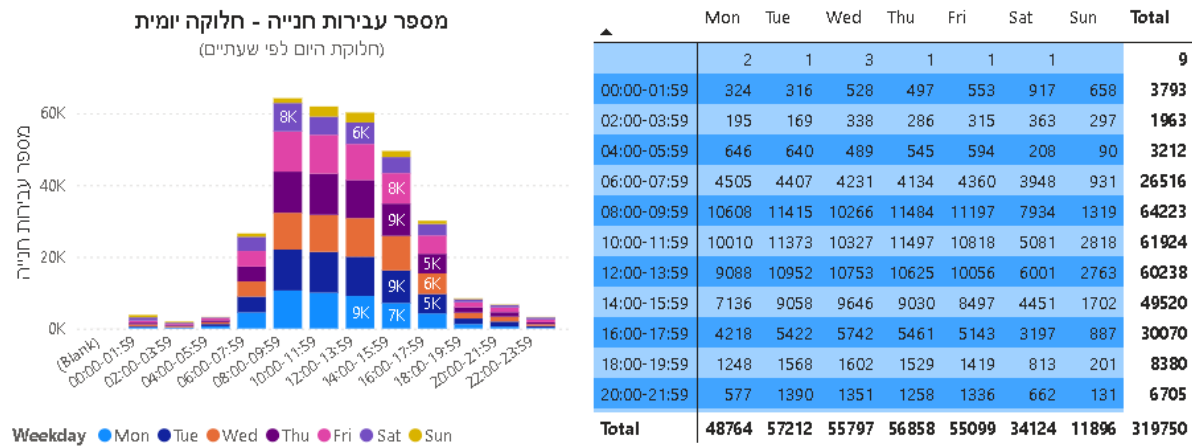
ביום חמישי הכי הרבה דוחות, ביום ראשון הכי פחות ובשבת אחרי ראשון (זה כנראה מספר על
 האוכלוסייה היהודית בניו יורק) ביום חמישי הכי הרבה – אולי כי לקראת סוף שבוע כולם מתאמצים
 לגמור עבודה (כולל אלו שבשישי ושבת לא עובדים)

מבחינת יום עבודה אין הבדל בין הרובעים, ביום שבת אופי העבודה בכל הרובעים הוא בערך 10%

בעיקרון בסטטן איילנד יש משמעותית פחות דוחות באופן גורף משאר שאר הרובעים – גם בשבת זה
 כך, ניתן לראות שמבחינת מחוזות אחרי סטטן, בברונקס יש פחות דוחות מאשר ב שאר הרובעים



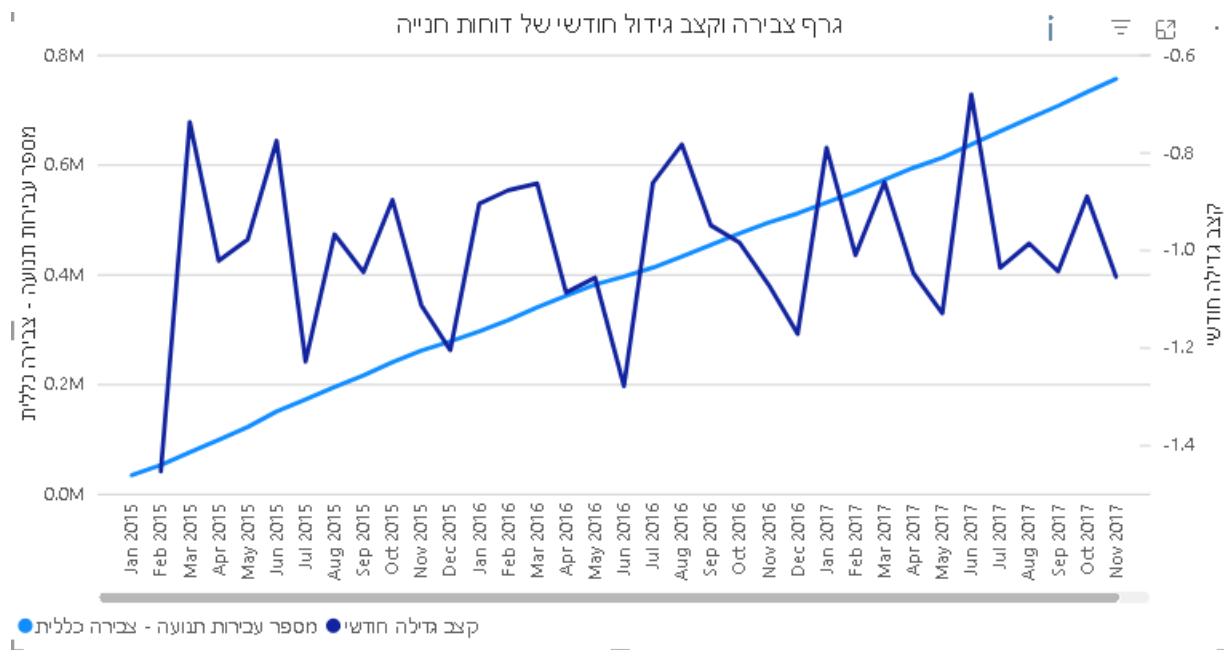
3.1.3 באילו שעות של היום, בפרקי זמן של שעתיים, יש יותר עבירות חניה? האם זה תלוי ברובע?



בין 0800 ל 1600 (שעות העבודה הקונבנציונליות) – יש יותר דוחות בעיקרון בין 1000-0800 (אולי זה קשור בהורים שלא למדו לשחרר ????)

המלצה : לבדוק סוגי דוחות ואזורים – יכל להיות שבניית מגרש חנייה ליד כל בתי הספר יכול להועיל כאן (אם אין עדיין)

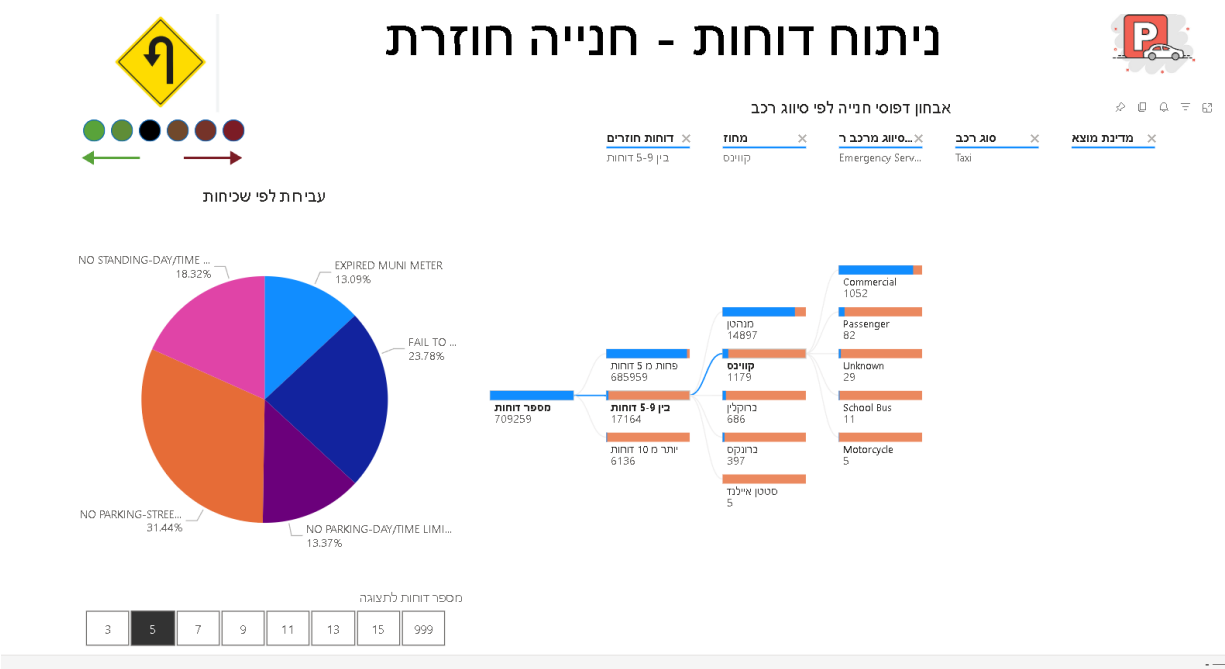
הציגו את מספר הדוחות בכל חודש לאורך השנים (2015-2017) ואת קצב הגידול החודשי



3.2 בנו ויזואליזציה המציגה את מספר המצטבר של הדוחות בכל חודש לאורך השנה YTD, לכל השנים



3.1.4 הציגו ויזואליזציה המראה כמה רכבים ביצעו יותר מ-10 עבירות חניה, כמה בין 5-ל-9 וכמה מתחת ל-5 עבירות



3.2 דוח

3.2.1 מהו סוג הרכב, המקבל הכי הרבה דוחות חניה ב NY?

SUV/SUBURBAN

כמות עבירות לפי סוג רכב

בחירת סוג רכב על פי סיווג כללי

All

מספר עבירות חניה	קוד סוג רכב	תיאור סוג רכב
251153	SUBN	SUV / Suburban
207236	4DSD	Four-Door Sedan
124578	VAN	Van Truck
66172	DELV	Delivery Truck
31372	SDN	Sedan
680511		

3.2.2 מהו גובה הקנס הממוצע לכל סוג רכב או לקבוצת סוגי רכב

גובה ההקנס הממוצע 75.83\$

סטטיסטיקות קנסות							
חלוקה לפי סוגי מרכב רכב							
סיווג כללי	מספר עבירות חנייה	ממוצע קנסות	ערך מינימום	ערך מקסימום	סה"כ הכנסות	סה"כ רווח	אחוז הרווח
Agricultural	80	81.69	35.00	165.00	6,535.00	6,135.00	93.88%
Commercial	244438	88.23	0.00	515.00	21,567,205.00	20,345,015.00	94.33%
Dealer	1	45.00	45.00	45.00	45.00	40.00	88.89%
Emergency Services	35	92.71	45.00	265.00	3,245.00	3,070.00	94.61%
For-Hire	4674	85.49	0.00	515.00	399,560.00	376,190.00	94.15%
Government	3	115.00	115.00	115.00	345.00	330.00	95.65%
Motorcycle	2242	76.21	0.00	115.00	170,855.00	159,645.00	93.44%
Passenger	513987	69.25	0.00	515.00	35,595,395.00	33,025,460.00	92.78%
Recreational Vehicle	99	73.94	35.00	115.00	7,320.00	6,825.00	93.24%
School Bus	3674	137.02	35.00	515.00	503,395.00	485,025.00	96.35%
Special Use	15	86.33	35.00	115.00	1,295.00	1,220.00	94.21%
Trailer	13	78.85	35.00	115.00	1,025.00	960.00	93.66%
Unknown	5014	89.88	0.00	515.00	450,635.00	425,565.00	94.44%
Watercraft	241	68.69	35.00	115.00	16,555.00	15,350.00	92.72%
Total	774516	75.82	0.00	515.00	58,723,410.00	54,850,830.00	93.41%

3.2 האם יש צבע רכב דומיננטי ?

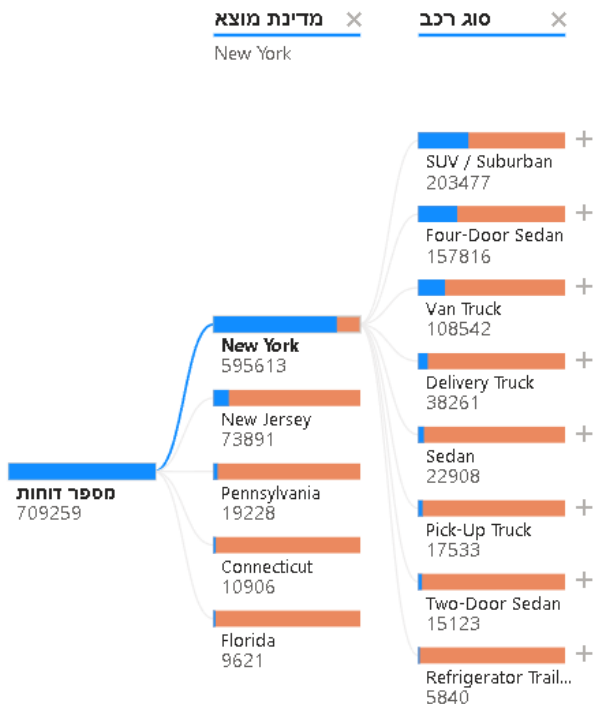
כן , לבן , אחר כך כחול ואחר כך שחור
 במקום השני קבוצת צבע לא מזוהה, כנראה זה לא שדה חובה

כמות עבירות לפי צבע

צבע	מספר עבירות חנייה
WHITE	234753
UNKNOWN	210722
BLUE	62667
BLACK	61888
GREY	56071
RED	44396
SILVER	25926
BROWN	25794
TAN	13176
YELLOW	10000

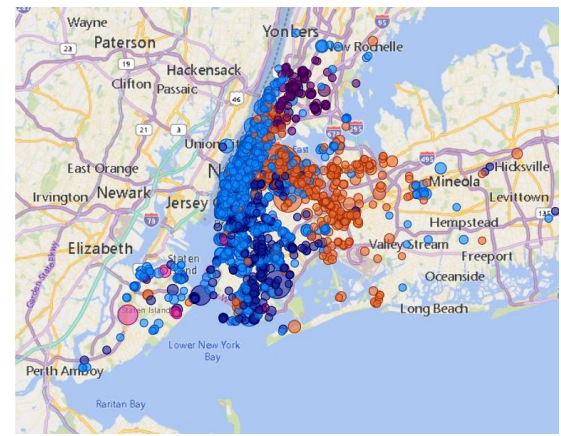
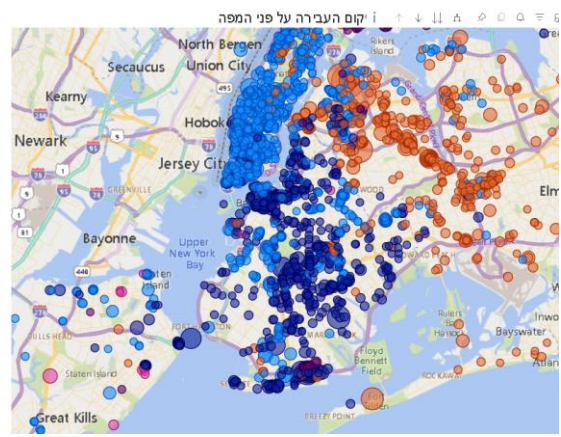
3.2.3 מאיזו מדינה State, באים רוב הרכבים המעורבים בעבירות חניה?

רוב הדוחות לרכבים מניו יורק , אחר כך ניו ג'רסי , אחר כך פנסילוניה



3.3 דוח 3

3.3.1 הציגו על מפה גיאוגרפית את עבירות החניה שבוצעו עם יכולת סינון של שעות היום והלילה. (הציעו חלוקה הגיונית של השעות ליום ולילה) (השתמשו בשדה Address עבור זיהוי המקומות על גבי המפה) (האם אפשר לזהות בצורה ברורה אזורים ו/או חלקי יום בעייתיים



על מנת לזהות מקומות על המפה נאלצתי לייצר שדה מחושב שמשורשר מהכתובת בית ומחוז

```
FullAddress =
CONCATENATE (
    CONCATENATE(VU_DimLocationCelan[HouseNumber] & " ", VU_DimLocationCelan[StreetName]),
    ", " & VU_DimLocationCelan[BoroughNameLookup] & ", New York, NY"
)
```

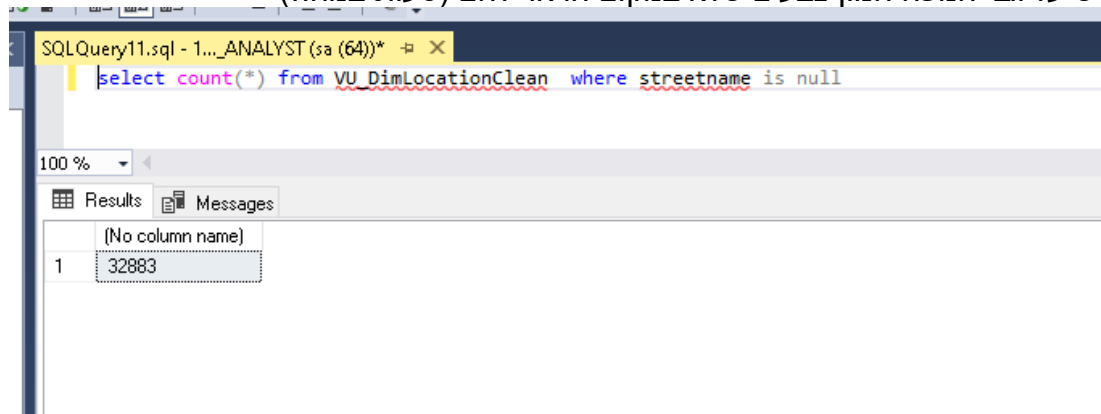
לא לכל הנקודות היה street , היו מקומות שהיה קוד רחוב.

בעיקרון אפשר לאחזר את זה באמצעות גישה לארכיב הרחובות של NY , זה קובץ טקסט בגודל MB 300

שצריך לקרוא בקוד די מסורבל, לא היה לי כבר זמן – אז וויתרתי – זה OVERKILL
זה אומר שחסרות על המפה 32883 נקודות

חוץ מזה, שהמון נקודות מדווחות לא נכון -, כמו שכבר רשמתי : כנראה פקח הכניס ידנית את המיקום, מספיק שטעה במחוז, או שה GPS שלו לא יהיה מכויל – והנקודה מתפספסת

יש על גבי המפה המון צבעים שלא במקום הראוי להם (טעות במחוז)



לגבי זמני לילה ויום : כתבתי קוד בפיתון (בעזרת GPT) שמייצר את כל שעות הלילה והיום של ניו יורק בין 2015-17

הקוד שמר ב CSV את כל זמני הזריחה והשקיעה , ייבאתי את הקובץ ל POWER BI
וכתבתי DAX שמייצר שדה IsDAy על סמך הנתונים הללו :

```
import pandas as pd
from astral.sun import sun
from astral import LocationInfo
from datetime import datetime, timedelta

# Define NYC Location
location = LocationInfo("New York", "USA", "America/New_York", 40.7128, -74.0060)

# Generate sunrise & sunset times for a full year
start_date = datetime(2015, 1, 1) # Change year if needed
end_date = datetime(2018, 1, 1)

date_list = []
sunrise_list = []
sunset_list = []
# Loop through each day in the range
current_date = start_date
while current_date < end_date:
    try:
        s = sun(location.observer, date=current_date, tzinfo=location.timezone)
        date_list.append(current_date.date())
        sunrise_list.append(s["sunrise"].strftime("%H:%M")) # Store sunrise
        sunset_list.append(s["sunset"].strftime("%H:%M")) # Store sunset
    except Exception as e:
        print(f"Error on {current_date}: {e}") # Debugging message
        sunrise_list.append(None)
        sunset_list.append(None)

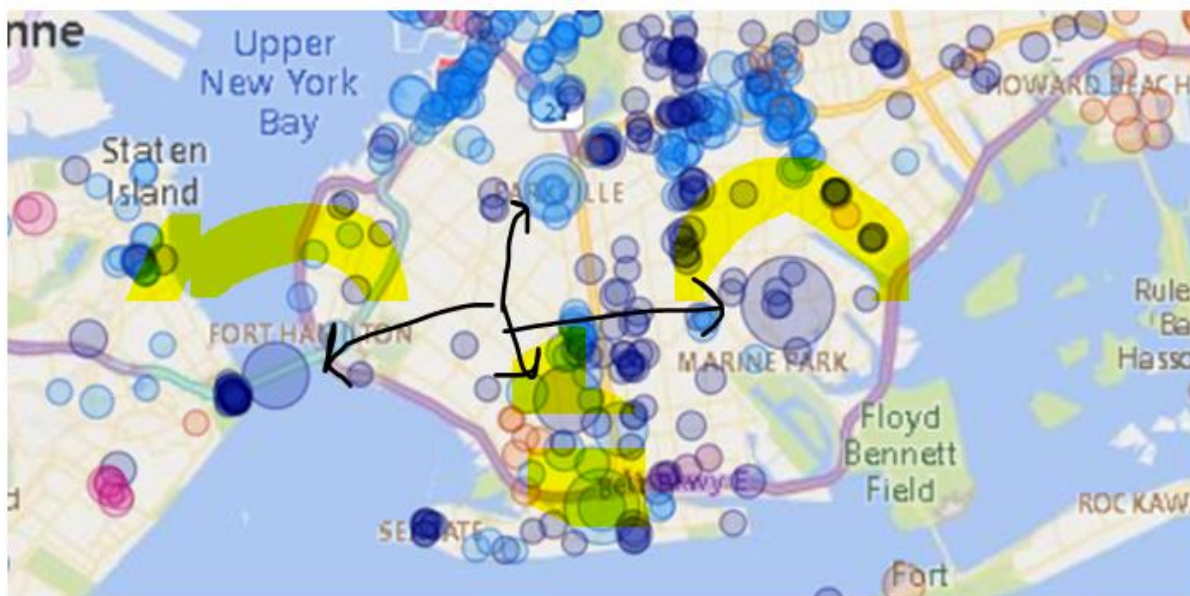
    current_date += timedelta(days=1)

# Create DataFrame
df = pd.DataFrame({"Date": date_list, "Sunrise": sunrise_list, "Sunset": sunset_list})
```

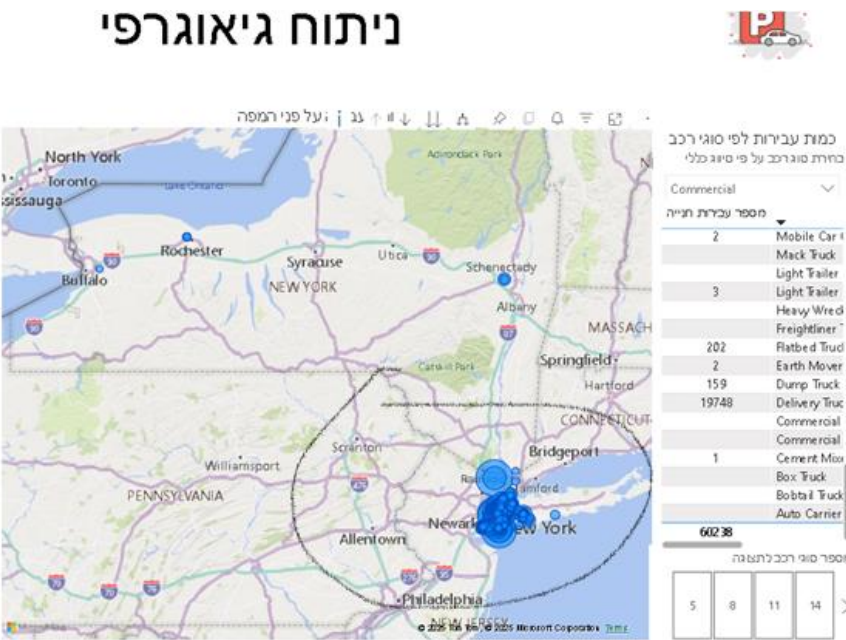
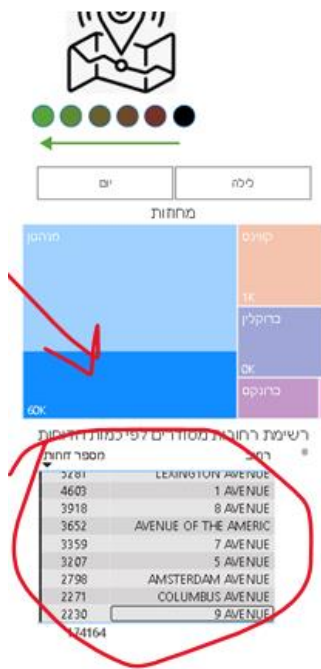
```
isDay =
VAR SunriseTime = LOOKUPVALUE(
'DimSunriseSunset_NYC2015To2017'[Sunrise],
'DimSunriseSunset_NYC2015To2017'[Date],
VU_FactTableFor2015To2017[ViolationDate] -- Match on Violation Date
)

VAR SunsetTime = LOOKUPVALUE(
'DimSunriseSunset_NYC2015To2017'[Sunset],
'DimSunriseSunset_NYC2015To2017'[Date],
VU_FactTableFor2015To2017[ViolationDate] -- Match on Violation Date
)

RETURN
IF(
    VU_FactTableFor2015To2017[ViolationTime] >= SunriseTime &&
    VU_FactTableFor2015To2017[ViolationTime] <= SunsetTime,
    "יום",
    "לילה"
)
```



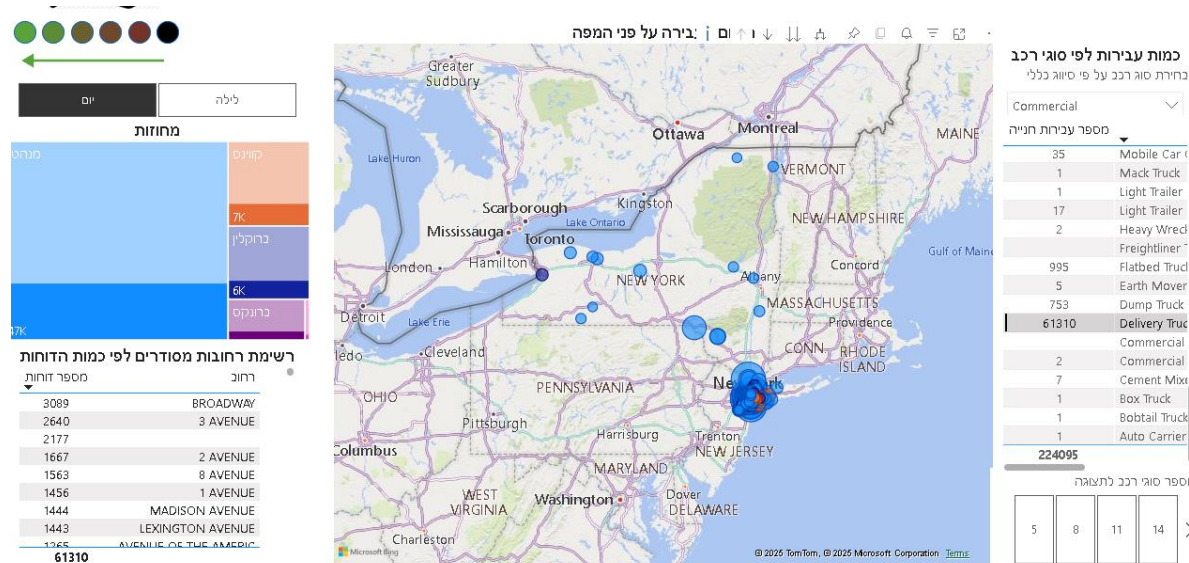
3.3.2 מה הן עשרת הרחובות במנהטן שיש בהם הכי הרבה דוחות חניה



שמקבלים בעלי
אלא כלל

רחוב	מספר דוחות	
AVENUE 3	8508	1
BROADWAY	8885	2
MADISON AVENUE	6217	3
AVENUE 2	5309	4
LEXINGTON AVENUE	5281	5
AVENUE 1	4603	6
AVENUE 8	3918	7
AVENUE OF THE AMERIC	3652	8
AVENUE 7	3359	9
5 AVENUE	3207	10

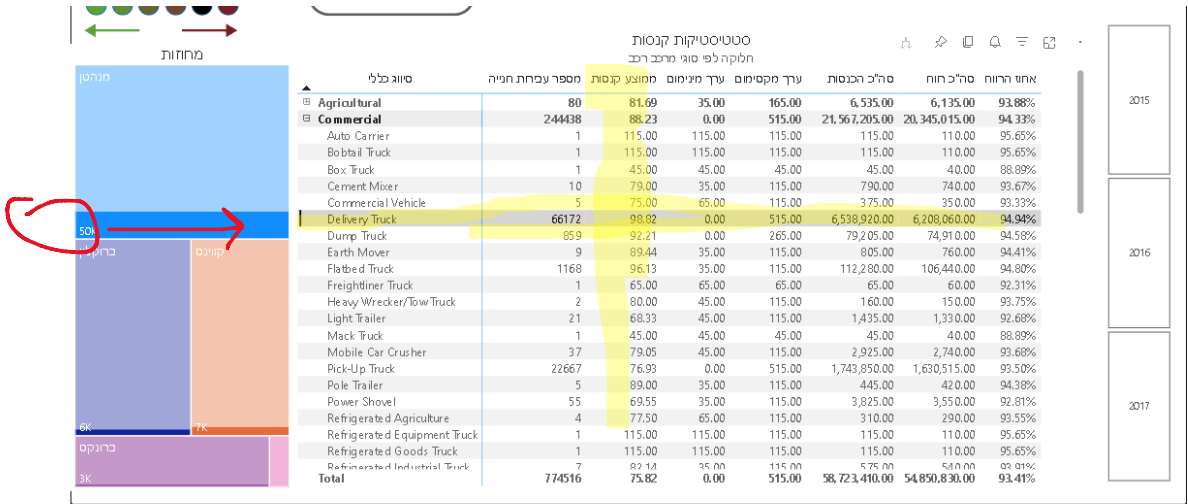
3.3 מה גובה הקנס הממוצע
רכב אלו? (לא רק העשרה העליונים
המשאיות



3.3.3 מה גובה הקנס הממוצע שמקבלים בעלי רכב אלו? (לא רק העשרה העליונים אלא כלל המשאיות)

הגובה הוא : \$98.82, לכל NY
אם נרצה לדעת מה הקנס הממוצע למחוזות בנפרד:

מנהטן	ברוקלין	ברונקס	קווינס	סטטן איילנד
\$100.94	\$98	\$93.45	\$87.05	\$72



```
PaymentAmountRecurrence =  
VAR IsManhattan = IF(VALUE(VU_FactViolationRecurrence[BoroughCode]) = 1, TRUE, FALSE)  
VAR FineAmount = LOOKUPVALUE(  
    DimViolationsFine[FineAmountUSD],  
    DimViolationsFine[ViolationKey], VU_FactViolationRecurrence[violationCode]  
)  
VAR FineAmountManhattan = LOOKUPVALUE(  
    DimViolationsFine[FineAmountUSDManhattan96],  
    DimViolationsFine[ViolationKey], VU_FactViolationRecurrence[violationCode]  
)  
RETURN  
IF(  
    ISBLANK(FineAmount) && ISBLANK(FineAmountManhattan),  
    0, -- Return 0 if no matching violationCode is found  
    IF(IsManhattan, FineAmountManhattan, FineAmount)  
)
```

3.4.1 מה ההכנסה, הרווח ואחוז הרווח (רווח חלקי הכנסות) של עיריית NY מאכיפת חוקי החנייה בכל שנה לכל סוכנות

2015

סטטיסטיקות קנסות חלוקה לפי סוכנות ממונה							
סוכנות ממונה	מספר עבירות חנייה	קנס מוצע	ערך מינום	ערך מקסימום	סה"כ הכנסות	סה"כ רווח	אחוז הרווח
BOARD OF ESTIMATE	145413	75.84	0.00	515.00	11,027,455.00	10,300,390.00	93.41%
DEPARTMENT OF BUSINESS SERVICES	2	90.00	65.00	115.00	180.00	170.00	94.44%
DEPARTMENT OF SANITATION	13	61.92	45.00	115.00	805.00	740.00	91.93%
FIRE DEPARTMENT	3370	48.97	0.00	180.00	165,035.00	148,185.00	89.79%
HOUSING AUTHORITY	1	115.00	115.00	115.00	115.00	110.00	95.65%
NYC TRANSIT AUTHORITY MANAGERS	11	86.36	60.00	115.00	950.00	895.00	94.21%
NYS PARKS POLICE	15	68.33	65.00	115.00	1,025.00	950.00	92.68%
OTHER/UNKNOWN AGENCIES	3	105.00	105.00	105.00	315.00	300.00	95.24%
PARKS DEPARTMENT	1357	90.77	0.00	180.00	123,180.00	116,395.00	94.49%
POLICE DEPARTMENT	97	83.30	45.00	165.00	8,080.00	7,595.00	94.00%
TRAFFIC	10556	101.34	0.00	515.00	1,069,735.00	1,016,955.00	95.07%
TRANSIT AUTHORITY	116606	73.32	35.00	515.00	8,549,160.00	7,966,130.00	93.18%
TRIBOROUGH BRIDGE AND TUNNEL POLICE	4	87.50	65.00	95.00	350.00	330.00	94.29%
Total	3	45.00	45.00	45.00	135.00	120.00	88.89%
Total	277451	75.50	0.00	515.00	20,946,520.00	19,559,265.00	93.38%

2016

חלוקה לפי סוכנות ממונה							
סוכנות ממונה	מספר עבירות חנייה	קנס מוצע	ערך מינום	ערך מקסימום	סה"כ הכנסות	סה"כ רווח	אחוז הרווח
BOARD OF ESTIMATE	5	97.00	65.00	115.00	485.00	460.00	94.85%
CON RAIL	13	65.38	45.00	115.00	850.00	785.00	92.35%
DEPARTMENT OF SANITATION	7219	48.67	0.00	165.00	351,365.00	315,270.00	89.73%
FIRE DEPARTMENT	12	110.00	95.00	115.00	1,320.00	1,260.00	95.45%
HOUSING AUTHORITY	4	91.25	60.00	115.00	365.00	345.00	94.52%
NYC TRANSIT AUTHORITY MANAGERS	13	66.15	0.00	115.00	860.00	795.00	92.44%
NYS COURT OFFICERS	4	100.00	95.00	115.00	400.00	380.00	95.00%
OTHER/UNKNOWN AGENCIES	418	84.58	0.00	180.00	35,355.00	33,265.00	94.09%
PARKS DEPARTMENT	143	88.22	35.00	180.00	12,615.00	11,900.00	94.33%
POLICE DEPARTMENT	19968	102.68	0.00	515.00	2,050,365.00	1,950,525.00	95.13%
PORT AUTHORITY	5	75.00	65.00	115.00	375.00	350.00	93.33%
TRAFFIC	205460	73.83	35.00	515.00	15,168,995.00	14,141,695.00	93.23%
TRANSIT AUTHORITY	3	96.67	60.00	115.00	290.00	275.00	94.83%
Total	233267	75.55	0.00	515.00	17,623,640.00	16,457,305.00	93.38%

סטטיסטיקות קנסות חלוקה לפי סוכנות ממונה							
סוכנות ממונה	מספר עבירות חנייה	קנס מוצע	ערך מינום	ערך מקסימום	סה"כ הכנסות	סה"כ רווח	אחוז הרווח
BOARD OF ESTIMATE	3	98.33	65.00	115.00	295.00	280.00	94.92%
DEPARTMENT OF BUSINESS SERVICES	3	83.33	60.00	95.00	250.00	235.00	94.00%
DEPARTMENT OF SANITATION	7803	49.61	0.00	265.00	387,110.00	348,095.00	89.92%
FIRE DEPARTMENT	11	107.73	95.00	115.00	1,185.00	1,130.00	95.36%
HOUSING AUTHORITY	15	78.67	60.00	95.00	1,180.00	1,105.00	93.64%
NYC TRANSIT AUTHORITY MANAGERS	1	115.00	115.00	115.00	115.00	110.00	95.65%
OTHER/UNKNOWN AGENCIES	395	80.42	0.00	165.00	31,765.00	29,790.00	93.78%
PARKING CONTROL UNIT	3	66.67	45.00	95.00	200.00	185.00	92.50%
PARKS DEPARTMENT	197	84.95	35.00	180.00	16,735.00	15,750.00	94.11%
POLICE DEPARTMENT	16189	101.63	0.00	515.00	1,645,215.00	1,564,270.00	95.08%
TAXI AND LIMOUSINE COMMISSION	1	115.00	115.00	115.00	115.00	110.00	95.65%
TRAFFIC	239170	75.55	0.00	515.00	18,068,380.00	16,872,530.00	93.38%
TRANSIT AUTHORITY	4	97.50	45.00	115.00	390.00	370.00	94.87%
TRIBOROUGH BRIDGE AND TUNNEL POLICE	3	105.00	95.00	115.00	315.00	300.00	95.24%
Total	263798	76.40	0.00	515.00	20,153,250.00	18,834,260.00	93.46%

2015

2016

2017

3.4.2 מה הרובע borough, הכי רווחי בעיר NYC?

מנהטן

להלן השוואת הרווח:

מנהטן	ברוקלין	ברונקס	קווינס	סטטן איילנד
\$ 27.2M	\$ 11.8M	\$ 5.5M	\$ 9.8M	\$ 0.56M

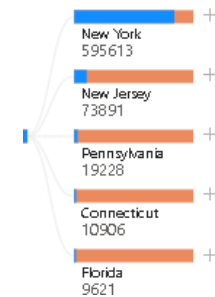
3.4.3 מה הם חמשת סוגי הקנס הרווחיים ביותר

No Standing – day/time limits
 No parking street cleaning
 Double Parking
 Fire Hidrant
 Fail To DSNLY MUNI METER RECPT

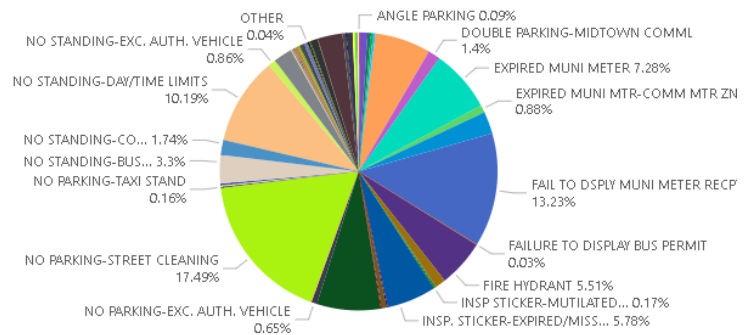
חוקה לפי סוגי חז							
סוגי חז	ס"כ רוח	ס"כ הכנסות	ערך מקסימום	ערך מינימום	קנס ממוצע	מספר עבירות חנייה	תוצר העבודה
NO STANDING-DAY/TIME LIMITS	95.65%	8,682,960.00	9,077,640.00	115.00	115.00	78936	
NO PARKING-STREET CLEANING	89.79%	5,957,620.00	6,635,010.00	65.00	45.00	135478	
DOUBLE PARKING	95.65%	5,562,590.00	5,815,435.00	115.00	115.00	50569	
FIRE HYDRANT	95.65%	4,697,220.00	4,910,730.00	115.00	115.00	42702	
FAIL TO DSNLY MUNI METER RECPT	88.61%	3,987,690.00	4,500,040.00	65.00	35.00	102470	
NO PARKING-DAY/TIME LIMITS	91.93%	3,281,855.00	3,569,925.00	65.00	60.00	57614	
NO STANDING-BUS STOP	95.65%	2,807,750.00	2,935,375.00	115.00	115.00	25525	
INSP. STICKER-EXPIRED/MISSING	92.31%	2,684,400.00	2,908,100.00	65.00	65.00	44740	
EXPIRED MUNI METER	89.29%	2,352,270.00	2,634,385.00	65.00	35.00	56423	
NO STANDING-EXC. TRUCK LOADING	94.74%	1,519,110.00	1,603,505.00	95.00	95.00	16879	
NO STANDING-COMM METER ZONE	95.65%	1,479,390.00	1,546,635.00	115.00	115.00	13449	
REG. STICKER-EXPIRED/MISSING	92.31%	1,340,940.00	1,452,685.00	65.00	65.00	22349	
FAIL TO DISP. MUNI METER RECPT	92.27%	1,183,890.00	1,283,025.00	65.00	35.00	19827	
DOUBLE PARKING-MIDTOWN COMM	95.65%	1,192,620.00	1,246,830.00	115.00	115.00	10842	
BIKE LANE	95.65%	742,940.00	776,710.00	115.00	115.00	6754	
SIDEWALK	95.65%	621,610.00	649,865.00	115.00	115.00	5651	
NO STANDING-EXC. AUTH. VEHICLE	94.74%	599,400.00	632,700.00	95.00	95.00	6660	
FRONT OR BACK PLATE MISSING	92.31%	567,720.00	615,030.00	65.00	65.00	9462	
EXPIRED MUNI MTR-COMM MTR ZN	92.30%	406,260.00	440,165.00	65.00	35.00	6781	
NO STOPPING-DAY/TIME LIMITS	95.65%	405,240.00	423,660.00	115.00	115.00	3684	
PLTRM LFTS LWRD POS COMM VEH	92.06%	378,800.00	411,470.00	65.00	45.00	6534	
OBSTRUCTING DRIVEWAY	94.74%	352,530.00	372,115.00	95.00	95.00	3917	
SAFETY ZONE	95.65%	332,640.00	347,760.00	115.00	115.00	3024	
COMM. PLATES UNREGISTERED VEHICLE	95.65%	326,700.00	341,550.00	115.00	115.00	3070	
Total	93.41%	54,850,830.00	58,723,410.00	515.00	0.00	774516	75.82

4. המלצות :

אני חושב שהעירייה חייבת לייצר שיח עם האזרחים של העיר.
 לא אנשים שמגיעים מבחוץ.
 כשאני מסתכל על מארג עבירות החנייה, אני רואה שפחות מ 20% מעבירות החנייה מתבצעות על ידי "תיירים" בעיר.



במבט על העבירות, אני רואה שהרבה מהעבירות היו יכולות להימנע הרבה עבירות של פתק לא מוצמד נכון , עבירות של חנייה ברחוב שמנקים אותו אלו העבירות שמכילות יותר מ 50% מכלל העבירות זה אומר שאין תקשורת בין האזרחים לבין העירייה, או שיש חוסר אכפתיות



אוכלוסייה נוספת שיש ללמד ולהשקיע בלימוד שלה כאן, היא אוכלוסיית הפקחים יש הרבה רמזים על חוסר מקצועיות בסיווג הדוחות, בזיהוי המקרים ואולי אף זלזול באזרחים אני לא מבין איך יכול להיות מצב שפקח רואה כרטיס חנייה ומתעלם ממנו בצורה כל כך גורפת, עד כדי שנותן דוח ומציין שהוא רואה כרטיס חנייה, אבל לא שמו אותו כמו שצריך

זה זלזול באוכלוסייה שאתה אמור לתת לה שירות

הסיווג הלקוי בסוג הרכב והטעויות במיקום, עוד יותר מחזקות ומגבשות בי את התודעה, שפקחים חייבים ללמוד להיות פקחים לפני שהם עובדים בזה
 אולי בכלל העברה של האחריות לחברה פרטית שמנסה להרוויח מזה, זאת הבעיה האמיתית.
 ניהול עיר זה לא רק עניין של רווח, אלא בראש ובראשונה דאגה לרווחת האוכלוסייה שלה.

לדעתי העירייה צריכה לנקות בכמה גישות:

1. הסברה – להסביר מתי מנקים את העיר, באילו ימים באיזה רחובות, ולהנגיש את המידע צריך ליישר קו לגבי מיקום כרטיס החנייה - חד משמעי
- אני חושב שדרך למנוע הישנות של התופעה היא לאפשר לפקחים להעביר קורסים (במימון העירייה) לאנשים להם נרשמים דוחות שכאלה
זה ייצר תקשורת בין אוכלוסיית הפקחים לאזרחים באזור השליטה שלהם ויוריד מוטיבציה לפקחים לרשום דוח שכזה
2. לפתוח חניונים נוספים במקומות בהם יש הצטברות של יותר מדי דוחות.
להבין את הבעייתיות ואולי לעזור בפתרונות חנייה נקודתיים עבור חניות בשעות בעייתיות בנקודות חמות
3. לגבי נהגים מסחריים – בייחוד כאלו שיש להם הרבה דוחות - יש לוודא שרציפי חנייה לשירות יהיו נגישים להם, ולהבין למה הם כן מסתכנים בקבלת דו"ח.
הרי ברור שהם יודעים שיקבלו ולמרות זאת הם חונים שלא כחוק
אולי אפשר יהיה לגבש פתרון שיעזור להם
4. להוציא את הפיקוח והאכיפה מגופים פרטיים שרק רוצים לעשות מכר רווח
הסיבה לתהליך הענישה אמורה להיות כדי לנוע את העבירה ולא על מנת להשריש את הבעייה ולהרוויח ממנה.
הגישה הזאת תבוא רק ארגונים ממשלתיים\עירוניים ינהלו את זה
5. הוספת אוטומציה לתהליך תמנע טעויות בהנפקת דוחות, ואפשר יהיה ללמוד את התהליך בצורה הוליסטית

הפעילות הרווחית ביותר היא במנהטן.
מצד שני המפה רוויה
אני הייתי מציע לעירייה להתחיל לאכוף מסביב למנהטן, לקחת את זה מתחת ידם של ארגונים פרטיים, ולאט לאט להתקדם למרכז התפוח
אולי מבחינת אכיפה היית ממליץ לנסות להגביל כניסה לעיר של נהגים שמועדים להרבה דוחות רכבים עם דוחות צריכים להיות מוזהרים ואולי אף לתת קנסות מוגברים לדוחות חוזרים

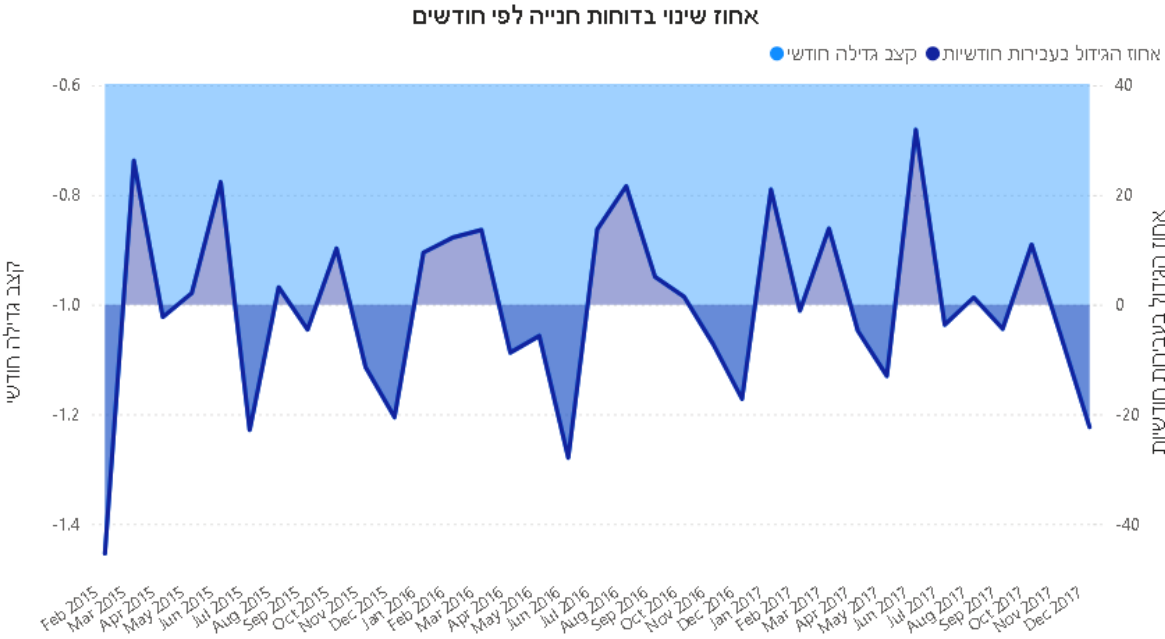
בעיקרון אני לא רואה שינויים מהותי.

אני רואה שיש ירידה בהכנסות וירידה גם ברווח – ביחס מתאים (מתון מאוד)

	סיווג כללי	אחוז הרווח	סה"כ רווח	סה"כ הכנסות	ערך מקסימום	ערך מינימום	קנס מוצע	מספר עבירות חנייה	סיווג כללי
2015	Total		19,559,265.00	20,946,520.00	515.00	0.00	75.50	277451	Total

2016	Total	233267	75.55	0.00	5 15.00	17,623,640.00	16,457,305.00	93.38%
2017	Total	263798	76.40	0.00	5 15.00	20,153,250.00	18,834,260.00	93.46%

יש יותר הכנסות ופחות רווח, מה שמספר שאולי יש יותר קנסות או
אנו גם רואים שבמבט על מדד הגידול בדוחות חנייה – הוא בירידה (כמון שכבר רשמתי)
וגם ראינו שקצב הגידול הינו שלילי



5. סיכום אישי

העבודה על הפרויקט היא מבורכת
נתקלתי בקשיים שלא צפיתי ולמדתי את הכלי ברמה טובה – לבד
מצד אחד התבאסתי שהדאטה גולמי וחסר ומצד שני היטבתי להסתדר ולמרות שהוצאתי על זה זמן, אני
מרגיש שהפתרון אליו הגעתי הוא מספק ודי מקצועי

נחשפתי בקורס לכלי מהנה שמספק לי הנאה מעבודה עם מידע.
אני איש של נתונים, וטכנולוגיות.
אין לי ספק – שהכלי (ואולי התפקיד ?) ימנף את עבודתי בעתיד

רציתי להרגיש את התחום והייתי סקפטי שיש לי מה ללמוד עוד
ובאמת השתכנעתי שאני יכול למצוא בו את עצמי.

כאיש טכני שיודע לתת נקודות שימוש נוספות לדאטה
מעבר למה שהכלי מספק – השמיים הם הגבול

תודה על האתגר

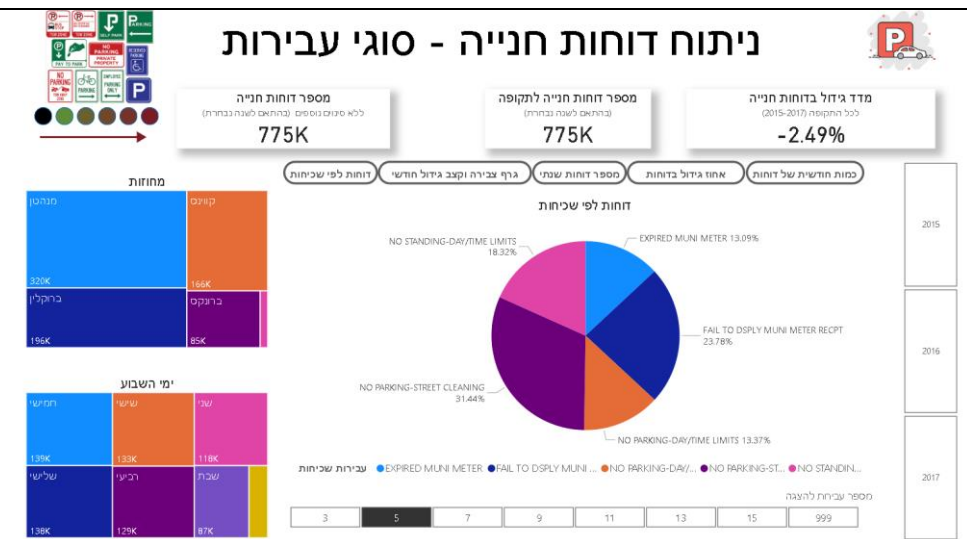
עידו

6. תיאור המערכת

מערכת הדוחות מכילה ממשק אליו ניתן להיכנס לדף הראשי בלבד :

<https://app.powerbi.com/groups/me/reports/323fdb2a-fd4f-4751-8a6d-99d9523cc096/8bf3bd0461ad5f4491c9?experience=power-bi>

<div><div></div><div>18/02/2025</div><div>פרויקט סיום קורס Business Analyst</div><div>בנושא : אכיפת דוחות חנייה בעיר ניו יורק בין השנים : 2015 – 2017</div><div>חוגש על ידי עידו ביסטרי ת"ז : 025704149</div><div> HORN TO START</div></div>	<p>יש ללחוץ על הדף כדי להתחיל</p>
--	---------------------------------------



הדוח מכיל גרפים מרכזיים (נגישים דרך הכפתורים שמעל לגרף המרכזי) ומערכת לסינון מידע לפי ימי השבוע, שנים ו/או מחוזות בניו יורק שבהם מתבצעת האכיפה של החנייה

ניתן לבחור כמה סוגי דוחות ייראו בגרף, בלחיצה על הכפתורים בתחתית הגרף 999 – מסמל כל הסוגים

המערכת מכילה 6 דוחות, שאפשר לנווט ביניהם בלחיצה על הכפתורים בצד שמאל:

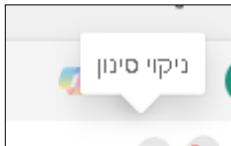


כל כפתור מנווט אותנו לדף אחר במערכת

בכל דו"ח – בצד ימין למעלה – יש תמרור חנייה אדום –



שמאפשר לנקות את הסינונים בדף



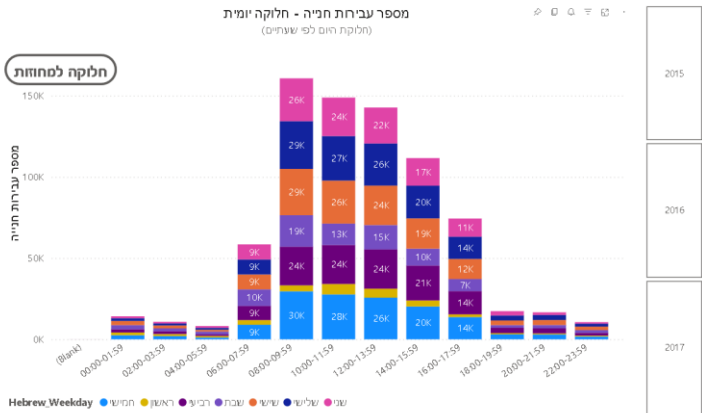
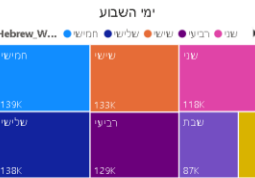
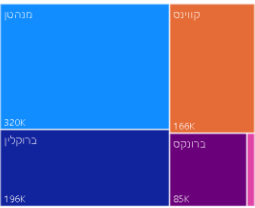
כמו שה ip toolt שלו מספר :

הדף הראשון מתמקד בהבנת המשקל של הסוגים השונים של דוחות החנייה ובסטטיסטיקה שלהם

הוא גם מכיל גרפים סיכומיים על משך חיי המערכת וההתנהלות שלה



ניתוח דוחות חנייה - ימי השבוע



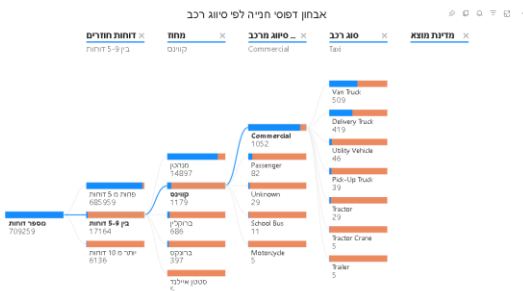
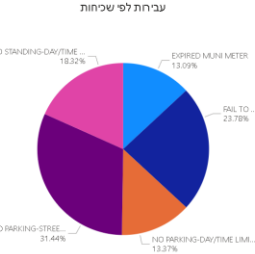
הדוח השני מתמקד בחלוקת מערכת הדוחות על פני היום – להבנת השימוש היומי של הנהגים החונים והבנת סף הלחץ – הזמן שבו יש הכי הרבה עבירות חנייה

גם כאן יש גרף מרכזי שניתן להתמקד בו משני מימדי מבט

מימד המחוזות
מימד היום



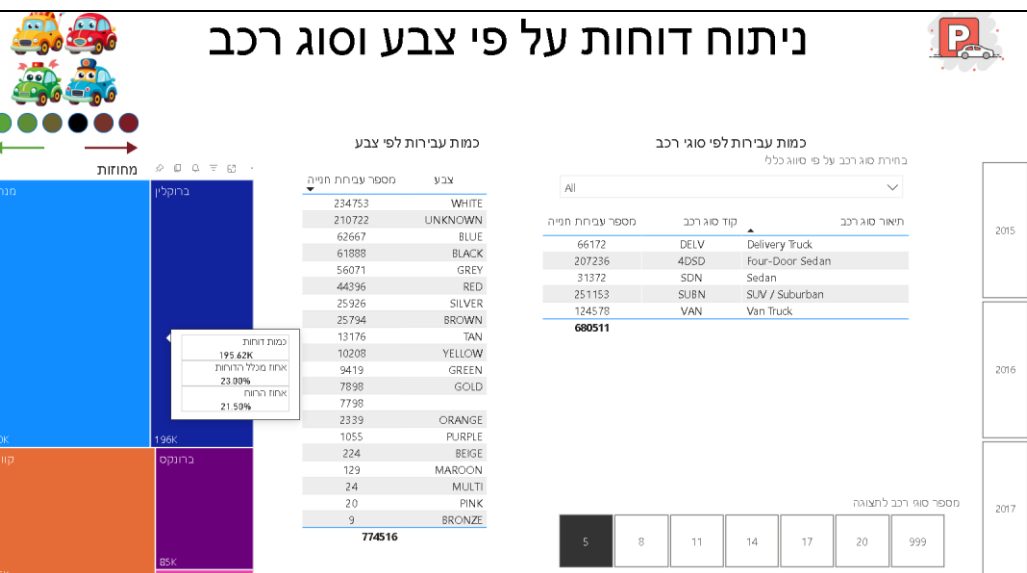
ניתוח דוחות - חנייה חוזרת



הדוח השלישי מתמקד בניתוח רב מימדי של המערכת : ניתן לשנות את נקודת ההתחלה של השאילתה ולחקור בה כל פרמטר

אפשר להסתכל על הדוח מנקודת מבט של נהגים עם דוחות חוזרים או להסתכל מנקודת מבט של סוג רכב

אנו מפנים כאן את נקודת המבט לפרמט נוסף : סוגי המכוניות שחונות כנגד החוק



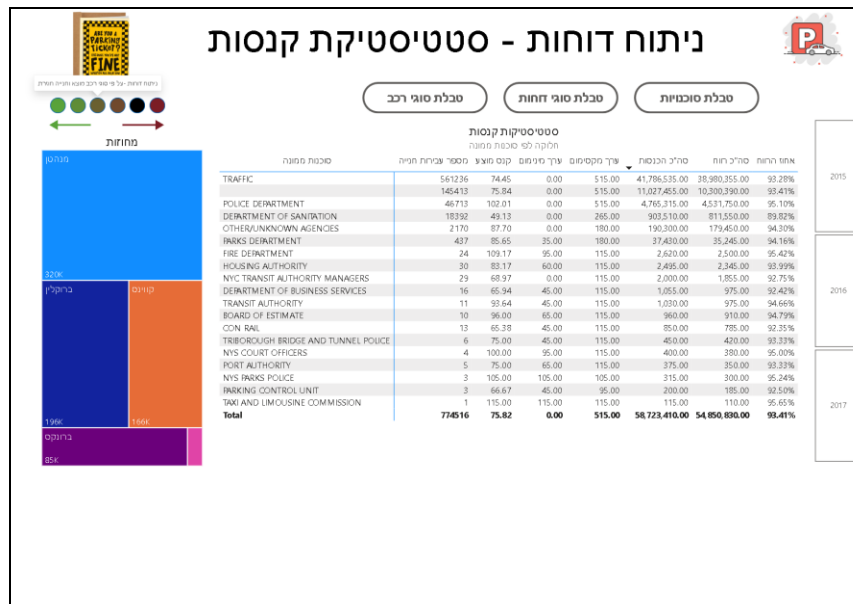
הדוח הרביעי נוגע בנקודת
הרכב כעיקרון

אנו מתחקים אחר סוגי הרכב
השונים והצבע

גם כאן אפשרות לבחור
במספר סוגי המכוניות
שחוקרים (בקוביות
שבתחתית הטבלה)

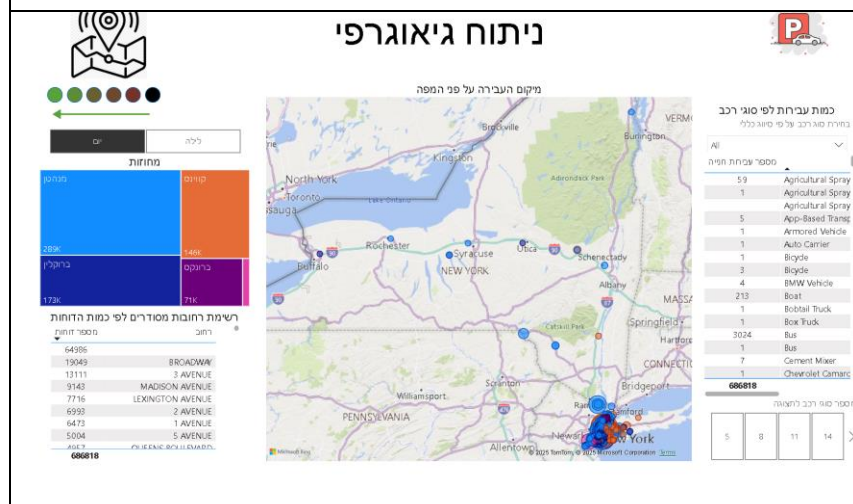
שימו לב שהצבע השחור מספר לנו כל הזמן את מיקומנו בזרימת הדוחות :





הדוח החמישי מביא אותנו
לסטטיסטיקות על הקנסות,
מנקודות מבט שונות :
הסוכנים , סוגי הדוחות , וסוגי
הרכב – כדי להגיע להבנה מי
מרוויח ומי ניזוק מהדוחות הכי
הרבה

שני הפילטרים הכי חשובים
מלווים אותנו לאורך כל
התהליך – כדי לעזור לנו
לחקור את הממדים : ממד
הזמן וממד המקום (המחוז)



הדוח האחרון מאפשר לנו
לפגוש פיסית את הדוחות
והמיקומים האמיתיים של
כבירות התנועה

כאן אנו שומטים את מיד הזמן ומתמקדים ברזולוציה יותר גבוהה במקום – ברמת רחובות ובסוגי הרכבים שעוברים על החוק