



אוניברסיטת בן-גוריון בנגב
Ben-Gurion University of the Negev

קורס ויזואליזציה של מידע 364-2-1721

ניתוח נתוני פשיעה



דניאל פיעניקה 336170808

תמיר אלון 301108874

איתן במנולקר 302780762

24.01.2022

תוכן

3.....	רקע והגדרת הבעיה
3.....	מבנה משטרת ישראל
4.....	תיאור הנתונים
4.....	נתוני פשיעה
5.....	נתוני יישובים
5.....	נתוני אבטלה
6.....	נתונים דמוגרפיים נוספים
6.....	נתוני קורדינאטות
6.....	הכנת הנתונים
7.....	דיון בעבודות קשורות לנתונים דומים
7.....	התפלגות עבירות
7.....	ניתוח לפי שנים
8.....	עבירות ביחס לממוצע ארצי
8.....	שיעור עליה/ירידה הגבוה ביותר לפי ביישובים
9.....	מצב הערים הגדולות
10.....	ניתוח תיאורטי של הויזואליזציות (מודל WWH)
10.....	דשבורד מצב פשיעה כללי
13.....	דשבורד מפה אינטקטיבית
15.....	שיעור אבטלה
16.....	צפיפות אוכלוסייה לק"מ
17.....	אחוז אוכלוסיה ערבית לקטגוריות פשיעה
18.....	מגמת כמות תיקים לפי רבעונים

רקע והגדרת הבעיה

בכל חברה מתוקנת ישנה שאיפה לחיים ציבוריים ואישיים בצורה תקינה, כך שכל אחד שבחברה יוכל להרגיש ביטחון אישי במרחב החיים שלו, לשם כך חברה מתקנת לעצמה חוקים ואוכפת אותם על ידי גופים המוסמכים לכך.

עם תקומת מדינת ישראל ב-1948, הוקם גוף משטרתי פורמלי שהיה בעברו חטיבה צבאית בארגון ההגנה, והוא הגוף האמון על אכיפה החוק במדינת ישראל.

את פרויקט בקורס "ויזואליזציה של מידע" בחרנו לבצע בתחום הפשיעה במדינת ישראל. תחום זה מרתק אותנו באופן אישי ואנו מאמינים שנוכל על סמך הצגת ויזואליזציות נכונות להציג תובנות משמעותיות.

נתוני תיקי המשטרה מפורסמים לציבור ועליהם פרויקט זה מתבסס. בנוסף ננסה אינטגרציית נתונים עם נתונים ממשלתיים נוספים (אחוזי אבטלה, מאזן דמוגרפי ביישוב, מאזן כלכלי ביישוב וכד') אשר לא מונגשים על ידי משטרת ישראל.

מבנה משטרת ישראל

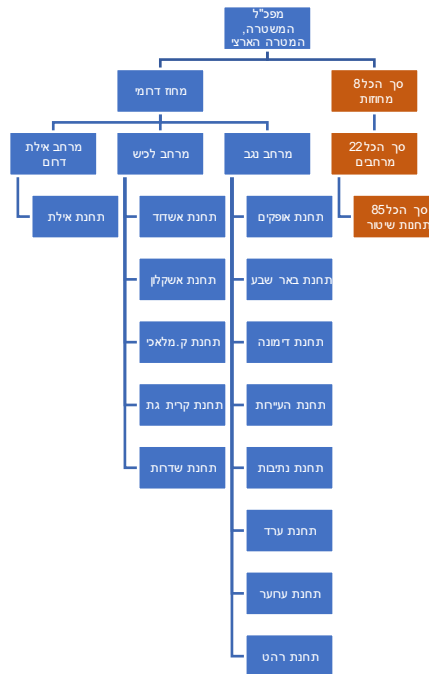
כדי להבין את נתוני המשטרה והפשיעה המוצגים בפרויקט זה נסביר מעט המבנה הנוכחי של משטרת ישראל על מנת שנוכל להבין את הנתונים השונים שיצגו הפרויקט זה.

משטרת ישראל הינו גוף היררכי, כל שכל תחנה כפופה למרחב, כל מרחב כפוף למחוז, והמחוזות השונים כפופים למטה הארצי של המשטרה בראשות מפכ"ל המשטרה.

נציין כי למשטרת ישראל גופים נוספים, ובהם לא נעסוק בפרויקט זה (משאבי אנוש, הדרכה וכד'), כיוון שבפרויקט זה אנו נעסוק בגופים האמונים על מיגור הפשיעה ובפתיחת תיקי משטרה.

בתרשים להלן ניתן להבין כיצד היררכית משטרת ישראל בנויה.

בתרשים מוצג היררכיית כלל התחנות והמרחבים במחוז דרום, וממנו ניתן להסיק את מבנה ההיררכיה לכלל המחוזות, המרחבים והתחנות השונות.



תיאור הנתונים

פאנל הנתונים הראשי של פרויקט זה הינו נתוני הפשיעה.

נתוני הפשיעה מונגשים על ידי משטרת ישראל, נתונים אלו הינם סיכומיים לפי רבעון, לכל תחנת משטרה ולפי סוג העבירה. הנתונים המוצגים בפרויקט זה הינם לשנים 2016-2020.

להלן פירוט נתוני הפשיעה:

נתוני פשיעה

נתוני הפשיעה נלקחו מאתר data.gov.il כאשר הטבלה מכילה את העמודות הבאות:

- רבעון ושנה – מימד זמן
- מחוז משטרה: דרום, חוף, ירושלים, מרכז, צפון, שי, ת"א, משמר הגבול
- מרחב משטרה: מרחב איילון דרום, מרחב איילון צפון, מרחב גליל צפון ועוד
- תחנת משטרה: תחנות המשטרה השונות בכל מרחב
- ישוב: ישוב אליו תחנת המשטרה משויכת
- קבוצה: קבוצת העבירה שבוצע כגון עבירות בטחון, רכוש, מוסר וכד'
- סוג: כל קבוצת עבירות מכילה פירוט נרחב יותר של העבירות.
- למשל בקבוצת העבירות יופיע "עבירות מרמה", ובסוג העבירות יופיע פירוט מעט רחב יותר: מרמה ועושיק, עבירות הנוגעות לכרטיס חיוב, זיוף מסמכים, סיוף כספים ועוד
- תיקים: מספר תיקים לפי הסדר ההיררכי המוצג (כמות תיקים שנפתחו ברבעון לתחנה מסוימת לפי סוג העבירה הרלוונטי)

לנתוני הפשיעה צירפנו נתונים נוספים שעשויים לשפוך אור נוסף על הקשרי הפשיעה השונים, גורמים קורלטיביים לפשיעה וכד' להלן פירוט הנתונים הנוספים שצירפנו לפרויקט זה:

נתוני יישובים

נתוני היישובים כוללים נתוני אוכלוסייה ומגזר. הנתונים נלקחו מהלשכה המרכזית לסטטיסטיקה. טבלת הנתונים מכילה את העמודות הבאות:

- שנה – מימד הזמן לפני שנים
- שם ישוב
- סמל ישוב: מספר מזהה ישוב
- תעתיק: שם הישוב באנגלית
- מחוז
- נפה
- אזור טבעי
- דת ביישוב – מספר דתות ביישוב (1-4)
- סה"כ אוכלוסייה – מס' התושבים בישוב
- סה"כ יהודים – מס' היהודים בישוב
- סה"כ ערבים – מס' ערבים בישוב
- שנת ייסוד – שנת ייסוד של היישוב

בנוסף, ראינו לנכון לצרף לפרויקט זה נתוני אבטלה, בתחושה הראשונית שלנו שיערנו שישנו קשר מסוים בין אבטלה לפשיעה. להלן פירוט הנתונים:

נתוני אבטלה

נתוני אבטלה נלקחו מאתר data.gov.il כאשר הטבלה מכילה את השדות הבאים:

- חודש ושנה – מימד זמן לפי חודש ושנה
- מחוז
- סמל עיר – מס' עיר
- שם עיר
- סוג עיר – יישוב או מועצה איזורית
- סה"כ מחפשי עבודה
- סה"כ מחפשי עבודה גברים
- סה"כ מחפשי עבודה נשים
- סה"כ מחפשי עבודה אקדמאים
- סה"כ מחפשי עבודה לא אקדמאים
- שיעור אבטלה מוערך

נתונים דמוגרפיים נוספים

הוצאנו נתונים כללים נוספים מאתר data.gov.il אשר יוכלו לסייע לנו בניתוח הנתונים. הטבלה מכילה

את האינדיקטורים הבאים:

- ישוב – שם הישוב
- כמות תושבים לק"מ
- סה"כ אוכלוסייה
- סה"כ אוכלוסיית נשים
- סה"כ אוכלוסיית גברים
- שיעור אוכלוסייה מגיל 0-4
- שיעור אוכלוסייה מגיל 5-9
- שיעור אוכלוסייה מגיל 10-14
- שיעור אוכלוסייה מגיל 15-19
- שיעור אוכלוסייה מגיל 20-29
- שיעור אוכלוסייה מגיל 30-44
- שיעור אוכלוסייה מגיל 45-59
- שיעור אוכלוסייה מגיל 60-64
- שיעור אוכלוסייה מגיל 65+
- שכר ממוצע
- שכר ממוצע גברים
- שכר ממוצע נשים
- שיעור השכירים אשר מרוויחים מעל שכר המינימום
- שכר ממוצע של עצמאים
- מדד אי שיוויון לכל יישוב (ג'יני)

נתוני קורדינאטות

על מנת שנוכל להשתמש בויזואליזציות מגוונות ונוחות, שיערנו שתצוגת נתונים על ני מפה תוכל לתת נדבך נוסף בפרויקט זה, על כן צירפנו לפרויקט נתוני GIS לכל יישוב.

הנתונים מכילים מיקום GIS (קורדינאטה X Y) לכל יישוב בישראל.

הנתונים נלקחו מאתר hub.arcgis.com.

הכנת הנתונים

על מנת שנוכל להשתמש במקורות נתונים הנ"ל, ביצענו הצלבות בין טבלת נתוני פשיעה ליתר הטבלאות. ייצרנו מפתח ייחודי המכיל קומבינציה של שני משתנים -שם היישוב ומימד הזמן. בנוסף, בוצע העמסות למשתנים כמותיים.

דיון בעבודות קשורות לנתונים דומים

אתר החדשות המפורסם- YNET ביצע ניתוח נתונים פשוט של נתונים פשיעה לשנים 2014-2019.

הכתבה פורסמה 11.06.20. קישור לכתבה: <https://www.ynet.co.il/articles/0,7340,L-5682378,00.html>

הנתונים הם נתוני טרור, קורונה, המשטרה טרם פרסמה את נתוני 2020 בתקופה זו ולכן הניתוח בוצע עד שנת 2019.

להלן מספר ניתוחים שבוצעו לגבי נתוני הפשיעה בישראל:

התפלגות עבירות

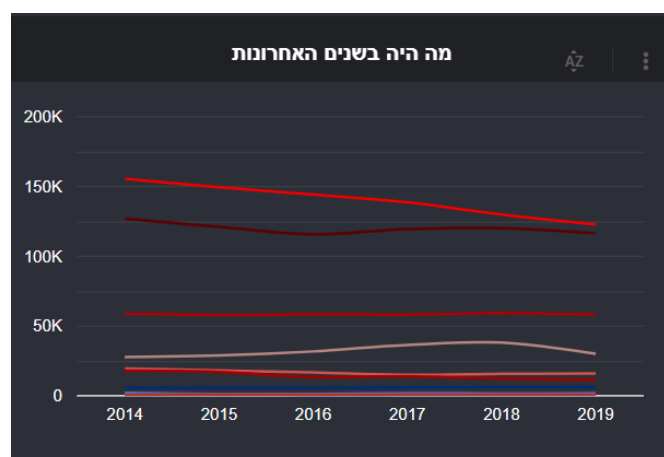
YNET בחרו להציג את התפלגות העבירות בPie-chart כאשר קיימת אפשרות פילטור לפי שנה, סוג עבירה ועיר. מויזואליזציה זו ניתן ללמוד כי שני שליש מהעבירות הינם עבירות של רכוש וסדר ציבורי.



לדעתנו בחירת הויזואליזציה אינה מיטבית כיוון שקשה להבחין ולהשוות בין גדלי קטגוריות העבירות, ובחירת הצבעים איננה טובה (הקטגוריות עבירות מין ו other אינם מכילות קונטרסט לצבע הרקע כמצופה).

ניתוח לפי שנים

YNET בחרו להציג את תמונת העבירות בLine chart. ניתן לראות כי קיימת ירידה מתונה במקרי העבירות



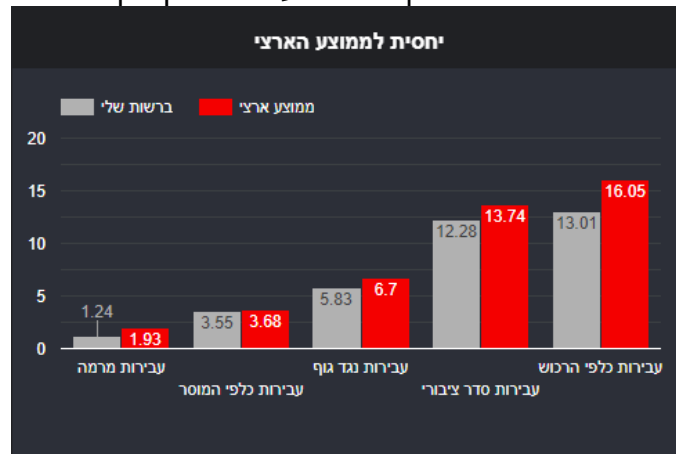
לאורך זמן כאשר מרבית הירידה היא בעבירות כלפי רכוש (150 אלף עבירות ב2014 לעומת 125 אלף עבירות ב2019).

גם כאן לדעתנו היה ניתן להציג את הנתונים בצורה טובה יותר, ניתן להבחין שקיימים המון נתונים שמצוגים כקו מגמה באזור ה0 ולא ניתן להבחין בין הקווים השונים.

עבירות ביחס לממוצע ארצי

YNET מציגים את חמשת קטגוריית העבירות הגבוהים ביותר באמצעות Bar plot עם שתי קטגוריות: ממוצע ארצי וברשות הנבחרת.

במקרה זה נבחרה לדעתנו ויזואליזציה טובה שכן הנתונים מוצגים בצורה בהירה ונוחה. נציין שלדעתנו חסרה האפשרות להרחבת כמות קטגוריות העבירות ובכך לקבל תמונה מלאה יותר.



שיעור עלייה/ירידה הגבוה ביותר לפי ביישובים

ניתוח מעניין שבוצע בכתבה זה מצביע על מגמות של יישובים עם העליות והירידות הגבוהות ביותר. לצד ניתוח זה, בוצע גם ניתוח של שיעור הירידה הגבוה ביותר ונמצא כי שיעור הירידה הגבוה ביותר. ויזואליזציה זו מוצגת כטבלה פשוטה שיכולה לספר תמונת מצב טובה במקרים מסוימים. החסרון המרכזי בויזואליזציה זו שלא צוינה צורת החישוב של אחוזי הגידול והירידה. בנוסף מוצגים שני פרמטרים, שגם אם כל אחד נכון כשלעצמו, לא ברור הצורך להציג אותם אחד לצד השני. כך למשל לא ברורה פשר העליה הגבוהה בפשיעה ביישוב היהודי היחיד בטבלה (ביתר עילית, עם עליה של 63%) לצד כמות התיקים לאלף נפש ב2019 (12 תיקים במקרה זה). שילוב הפרמטרים לא יצר אצלנו יותר בהירות, אלא יותר תחושה שאנחנו לא מבינים את מהות הנתונים

היישובים עם שיעור הירידה הגבוה ביותר 2019-2014

שיעור ירידה	תיקים לאלף נפש ב-2019	
-51%	9	שוהם
-49%	27	מועצה אזורית חוף השרון
-47%	19	מועצה אזורית גולן
-46%	15	יקנעם עילית
-46%	21	מבשרת ציון
-46%	16	קדימה צורן
-45%	23	תל מונד
-42%	34	סירת כרמל
-42%	31	מועצה אזורית באר טוביה
-42%	21	מעלה אדומים

היישובים עם שיעור העלייה הגבוה ביותר 2019-2014

שיעור עלייה	תיקים לאלף נפש ב-2019	
128%	48	ערערה בנגב
117%	48	טורעאן
75%	39	כפר כנא
63%	12	ביתר עילית
61%	35	כפר מנדא
58%	26	ירכא
56%	45	טמרה
55%	13	בית ג'ן
50%	30	אעבלין
45%	36	רהט
-15%	34	ממוצע ארצי

*ישובים עם יותר מ-10 אלף תושבים לפי נתוני למ"ס 2017

מצב הערים הגדולות

בערים הגדולות ניתן לראות ירידה בכמות התיקים. המשטרה טוענת כי בזכות מאבקה המתמשך נגד הפשיעה, שיעור הפשיעה הוא הנמוך ביותר מזה עשרות שנים. זאת בזכות פעולת אכיפה אפקטיבית ובזכות פעילות המשטרה על מגוון יחידותיה.

המצב בערים הגדולות מעל 100 אלף תושבים

שיעור שינוי 2019-2014	מספר תיקים לאלף נפש ב-2019	
-19%	30	ירושלים
-19%	67	תל אביב יפו
-27%	43	חיפה
-23%	28	ראשון לציון
-23%	29	פתח תקווה
-5%	33	אשדוד
-26%	34	נתניה
-4%	62	באר שבע
-29%	15	בני ברק
-22%	28	חולון
-28%	28	רמת גן
-7%	43	אשקלון
-22%	25	רחובות
-16%	44	בת ים
-17%	20	בית שמש
-15%	34	ממוצע ארצי

ניתוח תיאורטי של היזואליזציות (מודל WWH)

דשבורד מצב פשיעה כללי

What – המשטרה מתמודדת עם ניהול משאבים מוגבלים מצד אחד, ועם רצון להוריד את אחוזי הפשיעה הגבוהים מצד שני. המטרה האופטימלית לנתר משאבי שיטור בהתאם למצב הפשיעה, ובהתאם לסוג העבירות השונות.

דשבורד זה משלב ויזואליזציות בין מספר משתנים, בחלקי פירוט הגרפים נתאר את המשתנים הרלוונטיים

Why – על ידי הצגת נתוני התפלגות פשיעה וחריגות פשיעה יש לדרג הפיקודי הבכיר של המשטרה יכולת לשלט בניתור והורדת אחוזי הפשיעה במקומות מועדים.

על ידי הצגת הנתונים הנ"ל בצורה נוחה, הדרגים הפיקודיים יכולים לנתר את משאבי המשטרה וכמויות השוטרים לפי צרכים נקודתיים. למשל: להגביר ניידות סיור בתחנות משטרה שבהם יש אחוזי פשיעה גבוהים לצורך הורדת הפשיעה במקומות האלו.

How – לצורך כך יצרנו דשבורד אינטראקטיבי שיציג את כלל נתוני הפשיעה לפי סוג העבירות ולפי מרחב המשטרה המוניציפלי.

הדשבורד מורכב מארבעה חלקים מרכזיים:

- פילטר זמן:

נתוני הפשיעה המוצגים הינם סיכומיים לפי רבעונים, כך שעל ידי בחירת סימון תאריכים רלוונטיים ניתן יהיה לראות את כל נתוני הדשבורד לפי בחירת תאריכים באופן פשוט וידידותי.

פילטר הזמן בנוי משני חלקים- שנים ורבעונים, ללא תלות ביניהם. כך שניתן לבחור שנים מסוימות (לא בהכרח לפי סדר רציף) ורבעונים מסוימים ללא תלות בין השניים.

נתוני הזמן שמונגדים הינם קטגוראלים עם משמעות לסדר (Q1-2016, Q2-20016 וכך הלאה)

- טבלה מרכזית:

במרכז הדשבורד יהיה טבלה מטריונית כך שבשורות יהיו כלל המרחבים השונים של משטרת ישראל, בעמודות יהיו פירוט קבוצות העבירות. בתאי הטבלה יהיו נתונים ממרמלים לפי כמויות התיקים לפי פרמטר מקובל של כמות תיקים ל1000 תושבים.

הנרמול מתבסס על נתוני התיקים של משטרת ישראל והוא מתבצע באופן הבא:

$$N = \frac{n * 1000}{Total\ population}$$

כאשר: N- כמות התיקים ל1000 תושבים

n – כמות התיקים המקורית

Total population- מספר התושבים ביישוב הרלוונטי

הרקעים של תאי הטבלה יוצגו לפי סולם ערכים נומרי עולה, כך שהרקעים של ערכים נמוכים יהיו בצבע לבן/אדום בהיר, וערכים גבוהים יהיו צבועים בצבע אדום כהה לצורך זיהוי מהיר של אנומליות בנתונים.

גרף זה בנוי משני ערכים קטגוריאליים: מרחבים וקבוצות עבירות, ומשנה נומרי רציף אחד: כמות תיקים ל-1000 תושבים

- בחלק התחתון יהיה פירוט של הערים בכל מרחב. ברגע שתבצע לחיצה על תא מסוים בטבלה יוצג נתוני הערים המשויכות למרחב המסווג, והתפלגות כמות התיקים של התא הנבחר לפי היישובים שבמרחב.

את נתוני הערים בחרנו להציג ב-Barchart קלאסי, כיוון שלכל מרחב ישנם מספר ערים מצומצם שמשוך אליו התייחסנו לנתונים כאל נתונים קטגוריאליים שאותם נח לקרוא בתרשים עמודות פשוט. לכל עמודה הוספנו תיבת המכילה את כמות התיקים בכל עיר.

גרף זה יכול משתנה קטגוריאלי – שמות הערים, ומשתנה רציף – כמות התיקים שנפתחו בכל עיר בחלק הימני הצגנו את פירוט העבירות של כל קבוצת עבירות.

- לכל קבוצת עבירות יש מספר לא מבוטל של תתי עבירות בקבוצה. בחרנו להציג את הנתונים בתרשים עמודות אופקי, בסדר יורד של ערכים מהגבוהים לנמוכים.

בתצוגה זו שני יתרונות מרכזיים:

1. אפשרות ויזואלית מהירה 'לעבור' על כל הקטגוריה ולראות את כלל נתוני ההתפלגות (לעיתים יתכן מספר שאינו חריג ביחס ליתר העבירות, אך הוא משקף חריגות בגלל אופי העבירה, למשל עבירות חמורות של רצח נצפה למספרים נמוכים מיתר הקטגוריה). לשם כך ראינו חשיבות להציג לאנשי המקצוע את כלל הנתונים כדי שיוכלו לבחון בעצמם את סוגי התפלגות הנתונים.

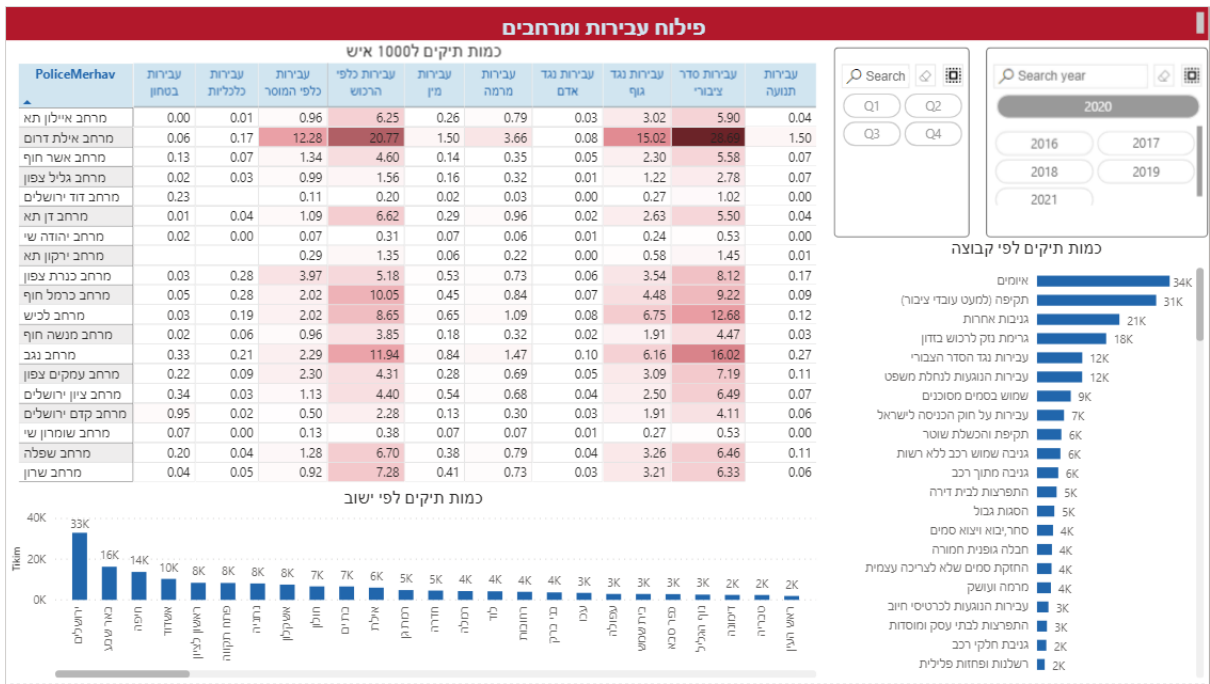
2. אפשרות מהירה לראות את סוגי העבירות הגבוהות ביותר והנמוכות ביותר, ובכך להבין האם המצב של המרחב הנבחר נחשב לחריג/ סביר וכד'.

גרף זה מכיל משנה קטגוריאלי – סוגי עבירות, ומשתנה רציף – כמות תיקים לפי כל סוג עבירה

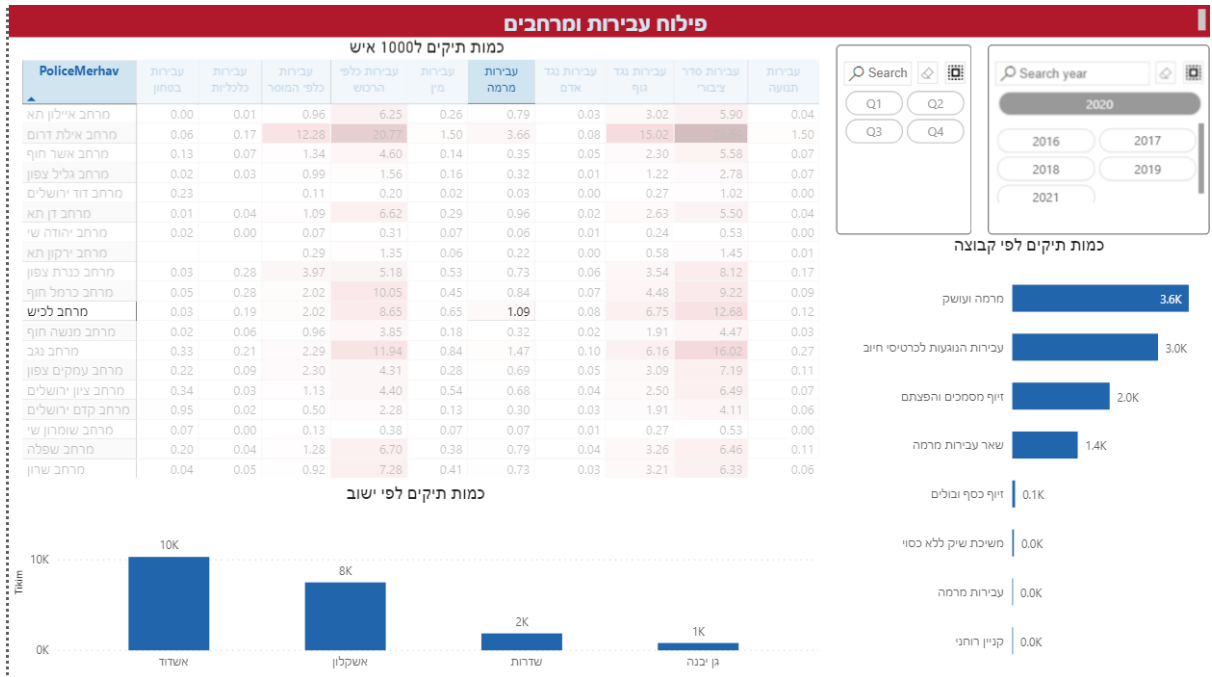
יש לציין שכשם שניתן לבחור תא בודד ולקבל ערכים אודותיו, כך ניתן לבחור עמודה שלמה ובך לראות את כל הנתונים אודות סוג העבירה הנבחרת בראיה ארצית. וכן ניתן לבחור שורה שלמה ובכך לראות את מצבו הכללי של המרחב, ולא רק בראיה מצומצמת של עבירה ספציפית.

לדשבורד זה בחרנו צבעים מתוך הפלטה הראשית של הצבעים שבחרנו להצגת נתונים אלו, תוך שימת לב מיוחדת להבלטת ערכים חריגים וניראות חדה של ניגודיות בין צבעים במקרה של בחירת רקע וגופן שונים.

דשבורד ראשי, ללא בחירת ערכים מסוימים



דשבורד במצב של בחירת ערכים מסוימים



דשבורד מפה אינטקראטיבית

What: דוברות משטרת ישראל נדרשה על ידי מפכ"ל המשטרה לספר נתוני פשיעה בצורה נוחה לציבור הרחב.

Why: משטרת ישראל מאמינה שעל ידי פרסום הנתונים גופים שונים יוכלו להירתם למשימה הלאומית בהורדת הפשיעה, למשל הגופים המוניציפליים, עמותות שונות ועוד.

בנוסף פרסום הנתונים ישווך כאקט שירותי לאזרח, ובכך אזרח פשוט יוכל לקבל החלטות שונות לגבי בהקשרי פשיעה, למשל: האם לגדל ילדים בישוב כזה או אחר? (לא נרצה לגדל ילדים בעיר עם פשיעה שמסוכנת לשלום הילדים, כגון עבירות מין), האם כדאי לגור ביישוב מסוים? (לא נרצה לגור ביישוב עם הרבה עבירות רכוש, סדר ציבורי נמוך וכד'), ועוד.

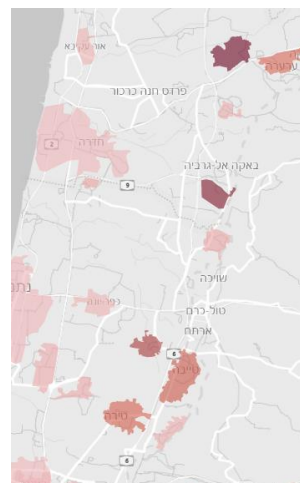
How: לצורך משימה זו בנינו דשבורד פשוט להבנה שכן הוא מיועד לכלל הציבור.

הדשבורד בנוי משלושה חלקים:

● מפה

החלק המרכזי בדשבורד הוא המפה, יתרונה של הצגת המפה הוא המעבר בקלות בין יישובים, אפשרות מהירה לעבור לנתונים של ערים סמוכות גאוגרפית (מה שלא אפשרי בתצוגה של מיון ערים לפי גודל אוכלוסייה, שם עיר וכד')

הערים השונות יצבעו בצבע כהה יותר ויותר ככל שרמת הפשיעה גבוהה יותר.
צבעית הערים תהיה לפי שטחה המוניציפלי של העיר, ובכך מאפשר בחירה נוחה לערים גדולות.
מעבר בין ערים מתבצע באופן פשוט למדי על ידי הזזת העכבר מעל המפה, כמו כן ניתן לבצע In/Out בכל חלק בטבלה.
צבעי המפה נבחרו לגוון אפור על מנת להבליט את האזורים הצבועים ברמות הפשיעה השונות,
לצורך הדגמה:



גרף זה בנוי מנתונים GIS של כל יישוב, ונתוני תיקים שנפתחו ביחס של 1000 תושבים

• חלק אינדקאטיבי –

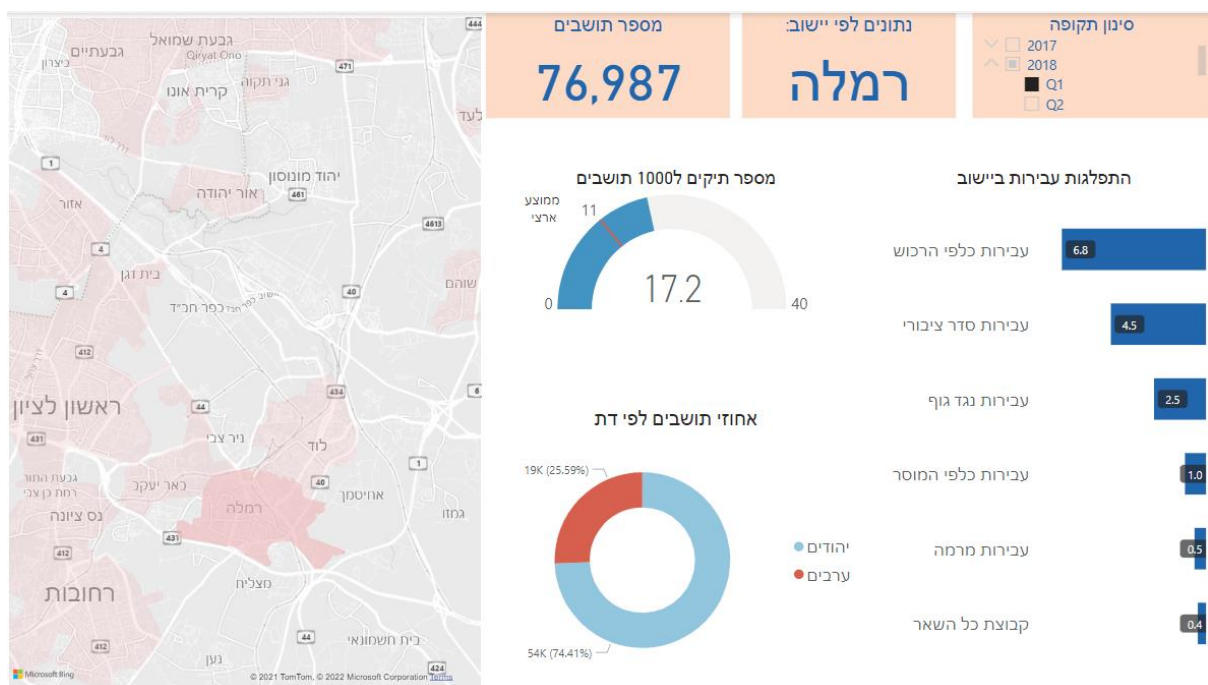
כיוון שמדובר בבחירה מתוך מפה, לעיתים יכולים להיות מצבי אי בהירות למשתמש בנוגע למקום הנבחר, לכן

ראינו לנכון לתת חיזוי למשתמש בנוגע לבחירת העיר.
 החיזוי נמצא בחלק העליון של הדשבורד, ומכיל את שם העיר שנבחרה ומספר התושבים בעיר.
 בחלק זה נמצא גם אפשרות בחירת תקופת הנתונים.

• מידע נוסף-

בעת בחירת עיר הצגנו נתונים נוספים רלוונטיים למשתמש:

- כמות התיקים ל1000 איש-
 בחרנו בתרשים Gauge שמראה באופן בהיר את כמות התיקים.
 יתרון משמעותי לחיזואליזציה זו היא ההשוואה לחס הממוצע הארצי.
 בנוסף בחירה זו מכניסה אלמנט מישחק חיובי בוכך יכולה להניע רשויות מוניציפאליות להניע למאמץ משולב יחד עם תחנת המשטרה המקומית להורדת הפשיעה מתחת לממוצע.
 גרף זה מכיל משתנה נומרי רציף של נתוני תיקים ל1000 תושבים.
 הגרף מציג באופן אוטומטי את מינימום ומקסימום שיעור התיקים שנפתחו בתקופה, ומציין את ממוצע פתיחת התיקים ל1000 תושבים לפי כלל היישובים באותה תקופה.
- התפלגות העבירות ביישוב-
 ראינו לנכון להציג למשתמשים את סוגי העבירות ביישוב ברמה ההיררכית הגבוהה שלהם, ובכך להציג פשטות ולהימנע מעומס נתונים שאינו נצרך ומצד שני לספק תמונת מצב כללית לגבי היישוב הנבחר.
 בחרנו בגרף עמודות אנכי המאפשר הצגת מידע באופן בהיר.
 גרף זה מכיל משתנה נומרי – כמות תיקים ל1000 תושבים, ומשתנה קטגוריאלי של קבוצות העבירות תושבים לפי דת-
- בחרנו להציג נתון זה על מנת לספק מידע דמוגרפי ביישוב.
 החיזואליזציה הנבחרת היא Pie-Ring שמאפשר קבלת מידע מידי ובהיר באופן בו ישנם מעט קטגוריות.
 בגרף מוצג משתנה קטגוריאלי עם חישובים על פני הקטגוריות הקיימות (אחוז וכמות נתונים).



שיעור אבטלה

What: מחלקת המחקר של המשטרה מנסה למצוא להבין את האם ישנם גורמים מסוימים שגורמים לפשיעה ובכך למגר את הפשיעה על ידי משאבי נוספים מעבר לשיטור מקומי (כגון השקעה בתשתיות, חינוך, רווחה ועוד)

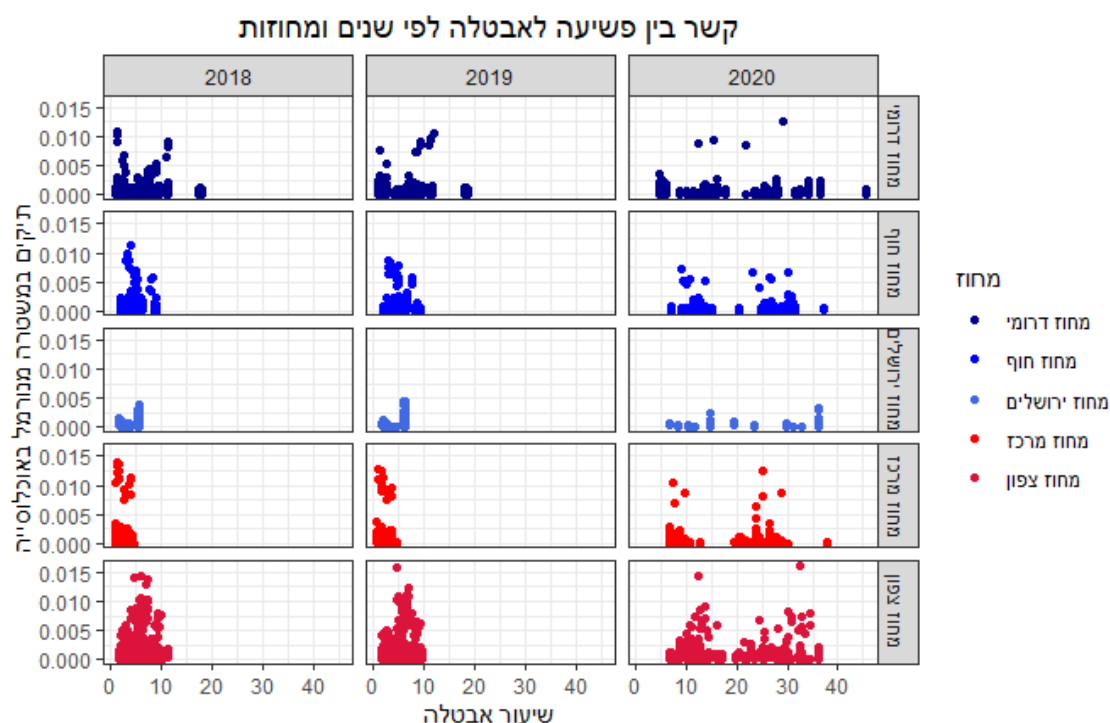
בגרף שיוצג ישנם שני משתנים רציפים: שיעור אוכלוסיה וכמות תיקים ממורמל ל1000 תושבים, ומשתנה קטגוריאל אחד – שנת הנתונים.

Why: בויזואליזציה זו נציג מגמות של נתוני אבטלה לצד נתוני פשיעה, אנו משערים כי נתוני אבטלה הינם נתונים שיכולים להשפיע באופן ישיר או עקיף על נתוני הפשיעה. כך שגם אם לא ימצא קשר ישיר, ההנחה היא שאלו נתונים משפיעים שיש להם חשיבות בתחום מחקר הפשיעה במשטרה.

How: בחרנו להראות גרף מגמה עם שני צירים (Scatter plot) בצורה מטריציונית, כך שנציג גרף מגמה כזה למחוז ולכל שנה מבין שלוש השנים האחרונות. בכל גרף בודד ציר הX מייצג את שיעור האבטלה וציר הY את מייצג את כמות התיקים ממורמל לפי גודל האוכלוסייה.

מימד המחוז נועד לצורך מחקר קשרים ייחודיים לפי מחוזות ספציפיים, ומימד השנה נועד לבחון האם יש שינוי תהליכי בקשר הקורלטיבי על פני השנים השונות.

להלן הויזואליזציה:



בשנים 2018-2019 שיעור האבטלה נמוך לעומת שנת 2020 בו שיעור האבטלה עולה עקב משבר הקורונה. ניתן לראות שאין קשר/קשר חלש חיובי בין שיעור אבטלה לפשיעה – גם בשנת 2020 ששיעור האבטלה עולה עדיין נראה שמס' התיקים נותר נמוך. ייתכן שישנו משתנה מאקרו-כלכלי/משתנה אחר שמשפיע על הפשיעה.

צפיפות אוכלוסייה לק"מ

What: מחלקת המחקר של המשטרה מנסה למצוא להבין את האם ישנם גורמים מסוימים שגורמים לפשיעה ובכך למגר את הפשיעה על ידי משאבי נוספים מעבר לשיטור מקומי (כגון השקעה בתשתיות, חינוך, רווחה ועוד)

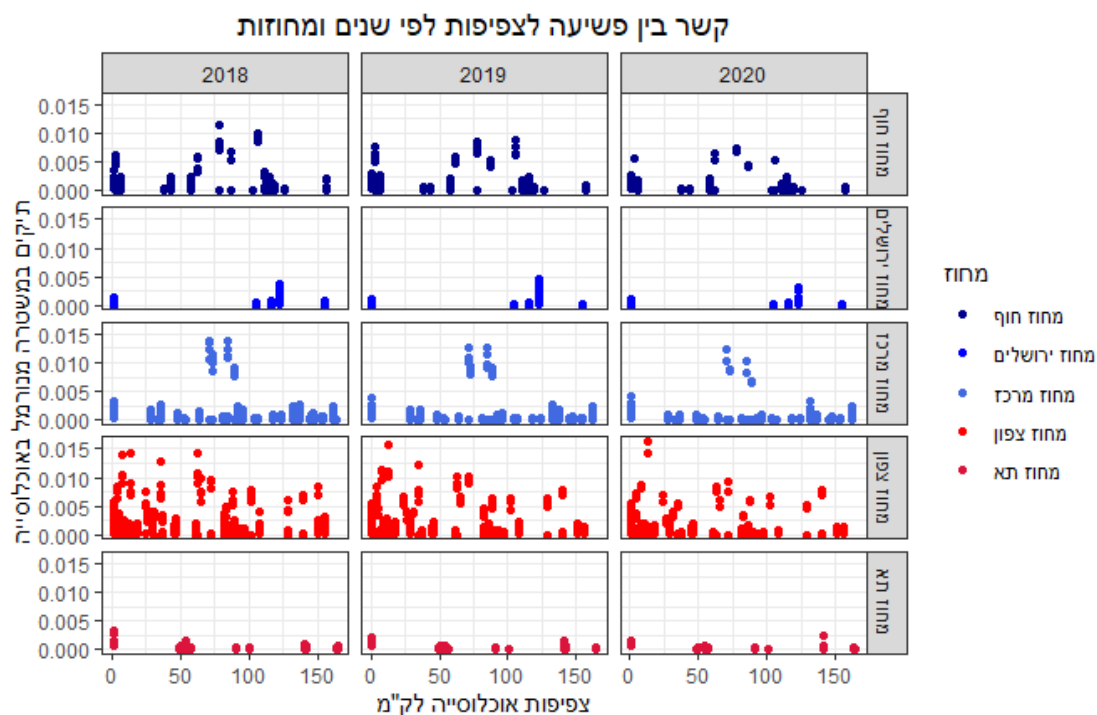
בגרף ישנם שני משתנים רציפים: שיעור צפיפות אוכלוסייה לק"מ רבוע וכמות תיקים מנורמל ל-1000 תושבים, ומשתנה קטגוריאלי אחד – שנת הנתונים.

Why: בויזואליזציה זו נרצה לבדוק אפשרות של הפשעת צפיפות אוכלוסייה על רמת הפשיעה, יתכן כי נתון זה מעיד על רמת עוני גבוהה יותר באזורים צפופים וכתוצאה מכך שיעורי פשיעה גבוהים.

How: גם כאן בחרנו להראות גרף מגמה עם 2 צירים (Scatter plot) בצורה מטריציאית, כך שנציג גרף מגמה כזה למחוז ולכל שנה מבין שלוש השנים האחרונות. בכל גרף בודד ציר ה-X מייצג את שיעור הצפיפות וציר ה-Y את מייצג את כמות התיקים מנורמל לפי גודל האוכלוסייה.

מימד המחוז נועד לצורך מחקר קשרים ייחודיים לפי מחוזות ספציפיים, ומימד השנה נועד לבחון האם יש שינוי תהליכי בקשר הקורלטיבי על פני השנים השונות.

להלן הויזואליזציה:



ניתן לראות שקיים קשר חלש בין צפיפות אוכלוסייה לפשיעה.

ספציפית במחוז צפון, ניתן לראות כי קיים קשר שלילי חלש בין המשתנים, ממצא שמפתיע אותנו מהסיבה שקיימים המון ישובים ערבים צפופים ואנו סברנו כי אלו נתונים קורלטיביים במידה מסוימת.

לעומת זאת, במחוז חוף, ניתן לראות כי קיים קשר חיובי חלש בין המשתנים.

אחוז אוכלוסיה ערבית לקטגוריות פשיעה

What – בתקופה האחרונה הגיעה פניה מצד תושבים רבים שהם מרגישים שישנה עליה בפשיעה הערבית, מחלקת המחקר של המשטרה נדרשה לאשש או להפריך טענה זו.

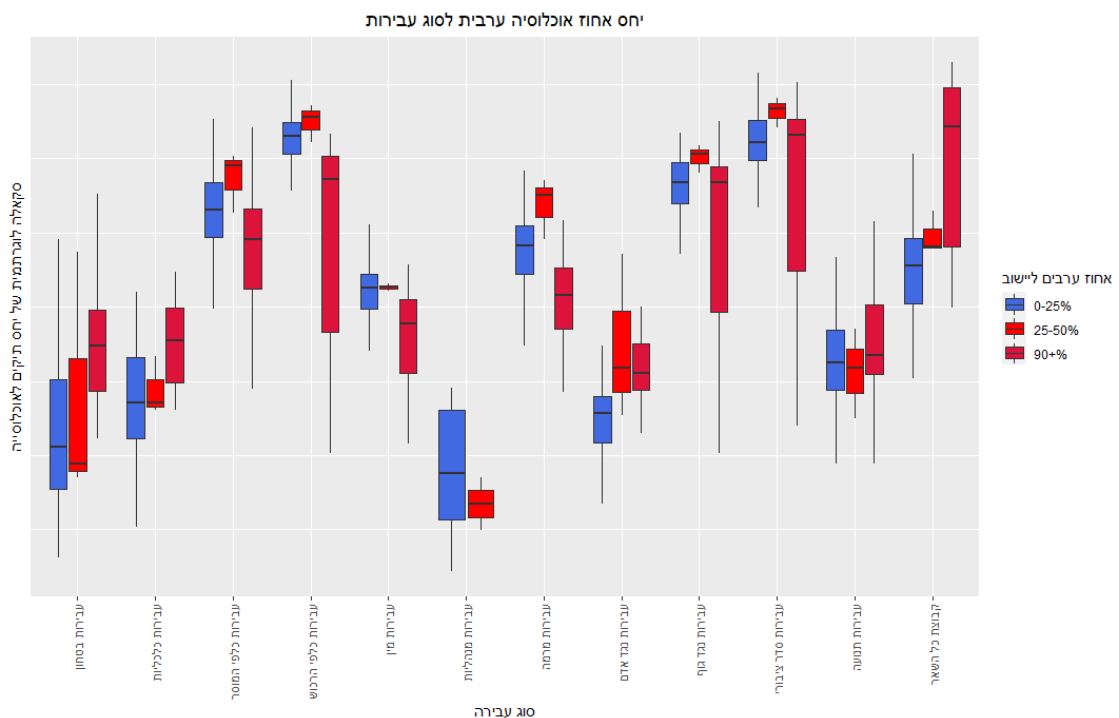
הויזואליזציה תכיל משתנה קטגוריאל ראשון של אחוז תושבים ערבים ביישוב: {0-25%, 25-50%, 90%+}, ומשתנה קטגוריאל נוסף של קבוצות העבירות.

Why – הבנה נכונה של קיום שיעורי פשיעה גבוהה באוכלוסיות מסוימות עשויה לסייע למשטרה במאמץ להפחתת הפשיעה במקומות מסוימים. עד כה משאבי המשטרה הושקעו לפי אזורים מסוימים, יתכן שאישוש טענה זו יוסיף נדבך חשוב בהבנת הפשיעה וסייע למשטרה לפנות לציבור הרחב לפי הקונטקסט התרבותי המקומי.

How – בחרנו להציג נתונים עבור כל סוג עבירה ולהראות Boxplot של משתנה קטגוריאל שמכיל אחוז של אוכלוסייה ערבית. ציר ה-X הוא הציר המתייחס לסוג עבירה. ציר ה-Y הוא הציר המתייחס ליחס תיקים לאוכלוסייה בסקאלה לגריתמית. הסיבה לבחירת ההצגה בסקאלה זו נובעת מכך שהיו כמה סוגי עבירות עם ערכים נמוכים מאוד (לדוגמא, עבירות ביטחון לעומת עבירות תנועה).

כיוון שאלו נתונים שמיועדים למחלקת המחקר ראינו לנכון להציג את הנתונים באופן זה, שכן אצד משתמש מורכב מאנשים בעלי הבנה מתקדמת בתצוגת נתונים מחקריים.

בנוסף, חילקנו את שיעור האוכלוסייה הערבית לפי קבוצות, זאת על מנת שנוכל להסיק על הבדלים מבחינת כמות תיקים לתושבים בהרכבים שונים של אוכלוסייה עבור סוגי עבירות שונות. להלן הויזואליזציה:



חשוב לציין כי אין פרמטר של 50%-90% כיוון שאין ישובים עם שיעור מעורבות כאלו, כשלעצמו זהו נתון מעניין מאוד.

ויזואליזציה זו עונה על השאלה: "האם עבור כל יישוב עם X% אוכלוסייה ערבית, אוכל לצפות ליותר או פחות

עבירות מסוג זה" על ידי תרשים Boxplot אנו יכולים להסיק האם יחס התיקים משתנה ביישובים שבהם יש אחוז גבוה יותר של אוכלוסיה ערבית.

למשל עבירות בטחון ועבירות כלליות- ניתן לראות שביישובים עם אחוז תושבים ערבים גבוה יותר יחס התיקים גבוה יותר. לעומת זאת, כשמסתכלים על העבירות מין ניתן לראות שאחוז תושבים ערבים גבוה לא משפיע על כמות התיקים, ואפילו להיפך – שיעור הפשיעה בעבירות אלו גבוה יותר כשאחוז התושבים הלא ערבים גבוה יותר.

מגמת כמות תיקים לפי רבעונים

What – לעדת הכספים של כנסת ישראל הגיעה טענה מעמותת "די למשטרה" כי תקציב המשטרה מנופח בהרבה ממה שנצטרך בפועל, לשם כך העמותה ביקשה להציג נתונים אודות כמות התיקים שנפתחו לפי תקופה.

הגרף מכיל נתונים זמן שמקרה זה הינם נתונים קטגוריאליים (Q1-2016, Q2-20016 וכך הלאה), ומשתנה נומרי רציף – כמות תיקים מנורמל ל1000 תושבים.

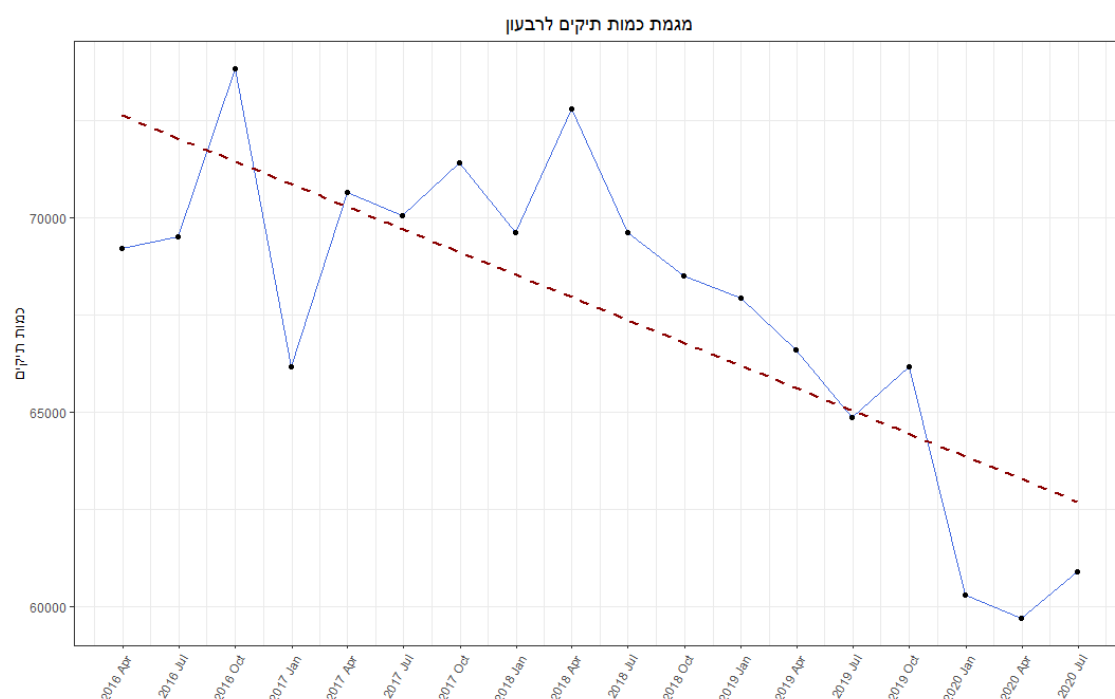
Why – לעמותה ישנה אג'נדה שריבוי משטרות לא מועיל למיגור הפשיעה, אלא יש להקצות את משאבי המשטרה למשאבי חינוך, רווחה והשערה ובכך למגר את הפשיעה.

How - מחלקת Datan של העמותה בחרה להציג תרשים מסוג Line Chart כדי לראות באופן פשוט ומיידי את כמויות פתיחות התיקים במשטרה. הויזואליזציה מורכבת משני צירים:

- ציר ה X מתייחס למימד הזמן, כלומר- רבעונים ושנים
- ציר ה Y מייצג את כמות התיקים שנפתחו באותה תקופה

ליזואליזציה נוסף קו מגמה שיכול ללמד בצורה אקסטרפולציונית על כך שכמות התיקים בשנה הבאה עתיד לרדת עוד יותר.

ליזואליזציה נוספו נקודות בולטות על מנת לסייע בהבנה מיידי ומדויקת של כמות התיקים שנפתחו באותה תקופה.



לפי קו הרגורסיה, ניתן לראות באופן מובהק את הירידה בכמות התיקים לפי רבעון. אם נסתכל על ה-Line chartn
אנו רואים ב2016 מגמת עלייה, בשנת 2017 יציבות ברמת התיקים ומ2018 ירידה חדה בכמות התיקים.
קשה למר מדוע הייתה ירידה דרסטית מהנתונים הקיימים ברשותינו. ייתכן ומדובר על הוספת תקציבים
למשטרה, הוספת משאבים וכו'.
בשנת 2020, ניתן לראות שפל בכמות התיקים עקב פרוץ הקורונה.