

ויזואליזציה של מידע



מגישים:

מאי חכים

גלעד ארז

[קישור לקוד בGithub](#)

[קישור לויזואליזציות בDrive](#)

(יש להוריד את קבצי הhtml ואז לפתוח בדפדפן)

תוכן

3רקע
3בסיס הנתונים
5תרשים 1: מספר התיקים שנפתחו לאורך הרבעונים
6תרשים 2: שינוי אחוזי בקבוצות-סוגי העבירות לאורך הזמן
7תרשים 3: סדר גודל סוגי העבירות הכלכליות לאורך הזמן
8תרשים 4: החלק היחסי (באחוזים) של Crime-Groups שונים לפי מחוזות
9תרשים 5: אחוז השינוי בקבוצות-סוגי הפשיעה השונים ברבעון של שומר החומות
10תרשים 6: אחוז השינוי בסוגי הפשיעה תחת עברות ביטחוניות ברבעון של שומר החומות
11תרשים 7: מפת מספר התיקים שנפתחו בכל מרחב
12תרשים 8: מפת פרופורציית Crime-Group בכל מרחב

רקע

משטרת ישראל אחראית על אכיפת החוק, שמירה על הסדר הציבורי ועל ביטחון הפנים של מדינת ישראל. משטרת ישראל היא כלל ארצית, והחלוקה לגזרות האחריות שלה נעשית לפי מחוזות. תחת כל מחוז ישנם מרחבים, ובכל מרחב מספר תחנות משטרה. מטרת הפרויקט היא לחקור את נתוני הפשיעה בישראל ברזולוציות שונות של סוגי פשיעה ותחומים גיאוגרפיים, על מנת למצוא מגמות מעניינות ולאפיין את הפשיעה לפי אזורים.

בסיס הנתונים

בסיס הנתונים שלנו הוא [מאגר נתוני הפשיעה של משטרת ישראל](#), אשר נבחר [ממאגר המידע הממשלתי של ישראל](#). מאגר זה מציג את מספר תיקי החקירה בהם יש חשד לביצוע עבירה פלילית, מקובצים לפי סוגי הפשיעה, רבעון ומיקום. הנתונים נאספו במשך חמש שנים (1-2018Q עד 1-2023Q). במאגר ישנה חלוקה גיאוגרפית לפי אזור סטטיסטי, קוד האזור, עיר, מחוז, מרחב ותחנת משטרה. המיקום של כל סוג תיק מסווג לפי המקום שבו בוצעה העבירה. התאמת הישוב לתחנה נעשה באופן אוטומטי לפי התחנה המפקחת על הישוב. העבירה עצמה מתוארת ע"י סוג העבירה וקבוצת העבירה, כאשר כל קבוצת עבירה מכילה מספר סוגי עבירות. זמן ביצוע העבירה מתואר ע"י שנה ורבעון. על מנת להוסיף ויזואליזציה מרחבית השתמשנו ב[מאגר הגבולות המשטרתיים של משטרת ישראל](#), אשר גם נבחר [ממאגר המידע הממשלתי של ישראל](#). מאגר זה מכיל שכבות גיאוגרפיות של הגבולות של משטרת ישראל- גזרות האחריות של התחנות, המרחבים והמחוזות השונים. השכבות הינן מסוג פוליון והקובץ מסוג GeoDataBase.

בבסיס הנתונים ישנם עשרה משתנים:

משתנה	תיאור
ArealAttributionType	האם התיק בטריטוריה ששויכה לאזור סטטיסטי
PoliceDistrict	מחוז משטרת
PoliceMerhav	מרחב משטרת
PoliceStation	תחנת משטרה
Settlement_Council	העיר בה בוצעה העבירה
StatArea	האזור הסטטיסטי אליו משויך התיק
Quarter	רבעון (למשל, 1Q-2018)
StatisticCrimeGroup	קבוצת עבירות (למשל, עבירות כלכליות)
StatisticCrimeType	סוג עבירה (למשל, עבירות פיסקליות)
TikimSum	סך התיקים

כיוון שאנחנו מעוניינים לסקור את סוגי העבירות לפי אזורים, החלטנו למחוק את הרשומות שאינן משויכות לאזור סטטיסטי. לאחר מכן מחקנו את העמודה המתאימה (ArealAttributionType) כיוון שכעת היה בה רק ערך אחד (תיקים באזור סטטיסטי).

צילום ממאגר הנתונים לאחר הורדת המשתנה ArealAttributionType:

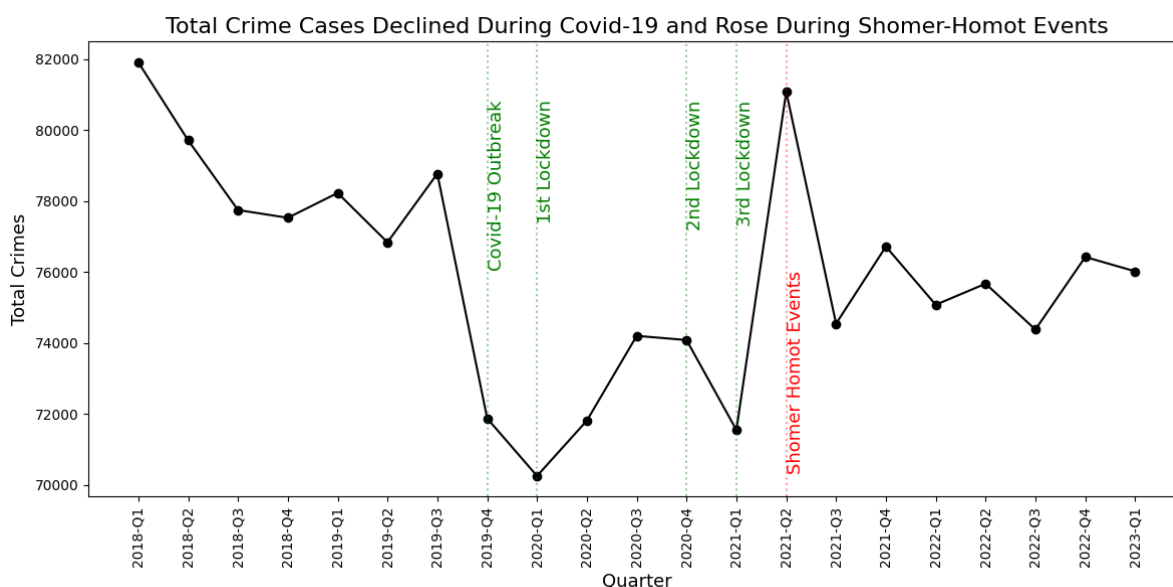
PoliceDistrict	PoliceMerhav	PoliceStation	Settlement_Council	StatArea	Quarter	StatisticCrimeGroup	StatisticCrimeType	TikimSum
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלכליות	עבירות פיסקליות	2
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי המוסר	קת סמים שלא לצריכה ...	3
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי המוסר	סחר, יבוא ויצוא סמים	3
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי המוסר	שימוש בסמים מסוכנים	44
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי הרכוש	גניבה מתוך רכב	5
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי הרכוש	גניבה ע"י עובדים וסוכנים	1
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי הרכוש	השימוש ברכב ללא רשות...	2
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי הרכוש	גניבות אחרות	74
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי הרכוש	גניבת אופניים	2
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי הרכוש	גרימת נזק לרכוש בדרך	12
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי הרכוש	החזקת רכוש גנוב	2
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי הרכוש	התפרצות לבית דירה	4
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי הרכוש	פרצות לבתי עסק ומוסדות...	2
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי הרכוש	כייסות	1
מחוז דרומי	מרחב אילת דרום	מרחב אילת דרום	אילת	26000011	2018-Q1	עבירות כלפי הרכוש	תקיפה לשם גניבה	1

רשומה לדוגמה: השורה הראשונה בצילום של מאגר הנתונים מתעדת שתי עבירות פיסקליות אשר נמצאות תחת קבוצת עבירות כלכליות, ברבעון הראשון של 2018 באילת, בתחנת מרחב אילת דרום, אשר שייכת למרחב אילת דרום, במחוז הדרומי.

על מנת להציג את הנתונים, משום שהם מוצגים בפורמט Long, השתמשנו באגרסיה על משתנים רלוונטיים ופעולת pivot. כך יצרנו טבלאות Short אינטואיטיביות יותר ונוחות לעבודה. לדוגמה, בטבלה המוצגת למטה ניתן לראות את מספר התיקים שנפתחו בסך הכל בכל מרחב משטרי לאורך השנים המתועדות במאגר (עמודת סה"כ), לצד מספר התיקים בכל קבוצת עבירה (Crime-Groups).

index	סה"כ	סעיפי הגדרה	עבירות בטחון	עבירות כלכליות	עבירות כלפי המוסר	עבירות כלפי הרכוש	עבירות מין
מרחב איילון החדש תא	177433	153	311	798	15146	69199	1005
מרחב אילת דרום	30192	29	44	117	3908	8445	104
מרחב אשר חוף	92761	38	167	593	7424	31889	292
מרחב גליל צפון	19604	15	167	169	1839	4654	82
מרחב דוד ירושלים	11118	11	791	34	670	1605	29
מרחב דן תא	77963	81	67	327	5108	29479	343
מרחב יהודה שי	12588	17	564	31	915	3239	81
מרחב ירקון תא	117660	127	75	491	8374	56779	518
מרחב כנרת צפון	51520	31	493	368	7191	12616	211

תרשים 1: מספר התיקים שנפתחו לאורך הרבעונים



מה:

נתונים שחישבנו מתוך הנתונים הגולמיים, המייצגים את מספר התיקים שנפתחו בכל רבעון בכל תחנות משטרת ישראל יחד.

משתנה כמותי (מספר התיקים) מסודר לפי משתנה סדרתי (הרבעון, קטגוריאלי אורדינלי ואינטרוואלי).

למה:

אנחנו מעוניינים לנתח מגמות בזמן, ולזהות חריגות במספרי התיקים. את שני אלה נצליב עם אירועים ארציים (ועולמיים) שחלו בתקופה שהנתונים מייצגים, בניסיון לחפש קשר בין האירועים הללו לרמת הפשיעה הכוללת בארץ.

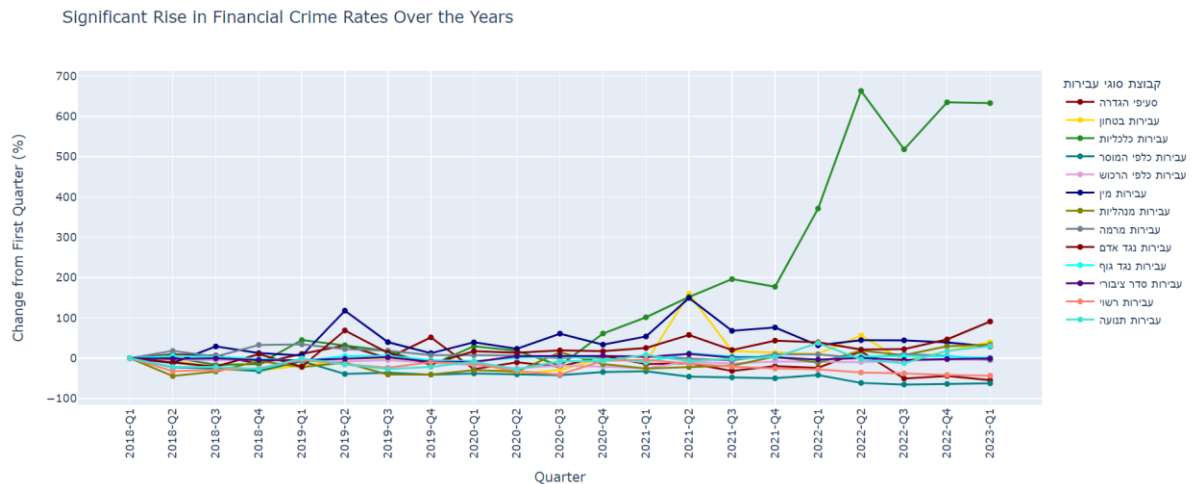
איך:

גרף קו.

על ידי אגרגציה של הנתונים לפי רבעונים וסכימה של מספר התיקים, קיבלנו טבלה בה מספר התיקים בכל רבעון. סידור הרבעונים באופן כרונולוגי ויצירת תרשים קו, מאפשרת התבוננות בשינויים ברמות הפשיעה לפי זמן.

על מנת להדגיש את זמן ההתרחשות של אירועים רלוונטיים, הוספנו קווים מנוקדים (הדפוס המנוקד נועד להבדלה מקו התרשים), והשתמשנו בצבעים אדום וירוק על מנת להפריד בין אירועי הקורונה שקרו בזמן ירידה בפשיעה, לאירוע שומר-חומות שקרה לצד עלייה בפשיעה.

תרשים 2: שינוי אחוזי בקבוצות-סוגי העבירות לאורך הזמן



מה:

נתונים שחישבנו מתוך הנתונים הגולמיים, המייצגים את אחוז השינוי במספר התיקים שנפתחו בכל Crime-Group בכל רבעון, ביחס לרבעון הראשון בנתונים. משתנה מספרי (אחוז השינוי) המחולק על פני 13 קבוצות עבירות (קטגוריאלית), מסודר לפי משתנה רבעון (קטגוריאלית אורדינלית ואינטרוולית).

למה:

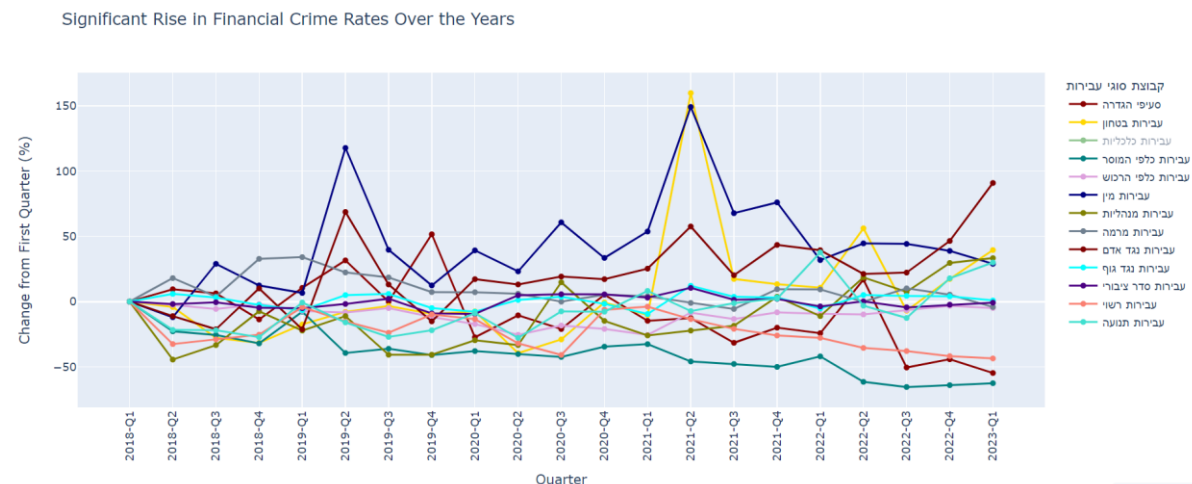
אנחנו מעוניינים להבין אילו סוגי פשיעה נהיו יותר או פחות נפוצים לאורך השנים. תובנות בנושא יוכלו לזהות תחומי פשיעה שמתגברים בארץ על מנת לרסן אותם (הגברת אכיפה, הרתעה, חקיקה...), או להצביע על שינויי מגמה חיוביים. ניתן לנתח את רמת ההצלחה של תוכניות לאומיות.

איך:

גרף קווים אינטראקטיבי.

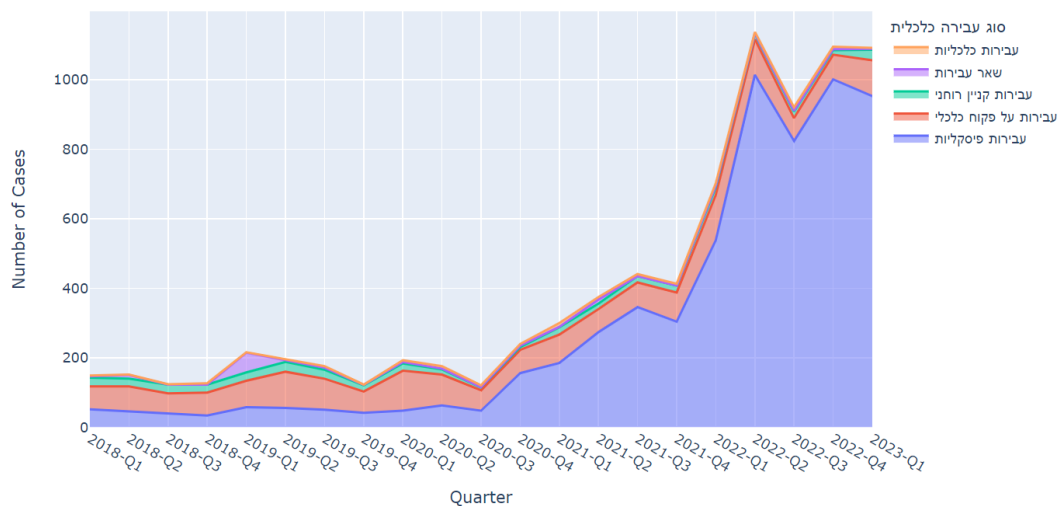
ביצענו אגרגציה של הנתונים לפי רבעונים ו-Crime-Group, וסכמנו את מספר התיקים. לאחר מכן חילקנו כל שורה (המייצגת קטגוריה = קו בתרשים) במספר התיקים שנרשמו ברבעון הראשון, וכפלנו ב-100 כדי להמיר מיחס לאחוזים. קיבלנו טבלה בה מוצג אחוז השינוי בכל קבוצת פשיעה בכל רבעון. סידור הרבעונים באופן כרונולוגי ויצירת תרשים קווים, מאפשרת התבוננות בשינויים וגם השוואה בין המגמות השונות שחוו קבוצות סוגי פשיעה.

צביעה של כל קו בצבע שונה מאפשרת הבדלה נוחה בין הקווים. משום שאין פלטות צבע שמאפשרות הבדלים בולטים בין כל 13 הקבוצות, יש דמיון בצבעים של חלק מהקווים - בעיה שנפתרת על ידי שימוש בתרשים אינטראקטיבי בו אפשר ללחוץ על קבוצה ולהעלים / להחזיר אותה לתרשים. גרף זה מאפשר גם משחק עם הציר האנכי, שמתאים את עצמו לקווים המופיעים. למשל - הסתרה של "עבירות כלכליות" מעדכנת את התרשים באופן הבא:



תרשים 3: סדר גודל סוגי העבירות הכלכליות לאורך הזמן

Rise in Financial Crimes Is Primarily Driven by Fiscal Offenses



מה:

נתונים שחישבנו מתוך הנתונים הגולמיים, המייצגים את מספרי התיקים שנפתחו מכל סוג-עבירה תחת קבוצת העבירות הכלכליות. משתנה מספרי (מספר התיקים) המחולק על פני 5 קבוצות (סוג העבירה הכלכלית, קטגוריאלית), ומסודר לפי רבעון (קטגוריאלית אורדינלי ואינטרוואלי).

למה:

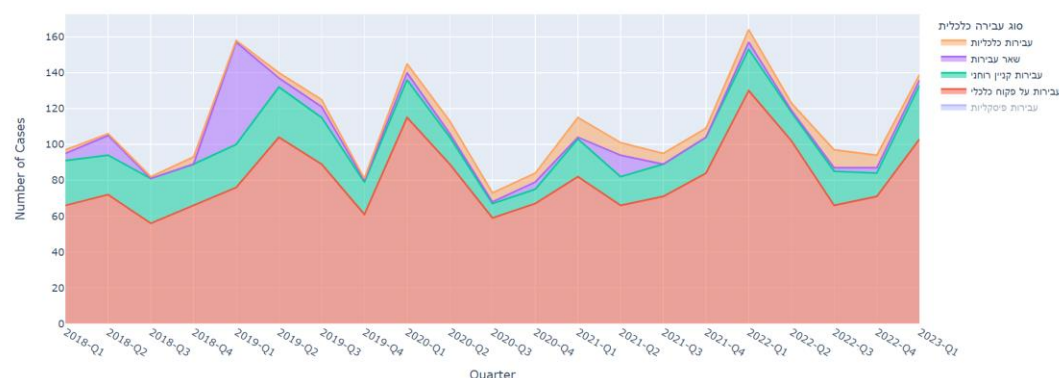
בגרף מספר 2 ראינו שמספר העבירות הכלכליות לרבעון עלה יותר מפי 6 (!) במהלך המעקב. על מנת להבין מה מתרחש בתוך קבוצת-סוגי העבירות האלה (Crime-Group), ולגלות מה מביא לעליה, יש לפרק את הקבוצה למרכיביה (Crime-Types). ניתוח זה יאפשר למקבלי ההחלטות למקד את הטיפול בסוגי העבירות הרלוונטיים.

איך:

גרף Stacked Area אינטראקטיבי.

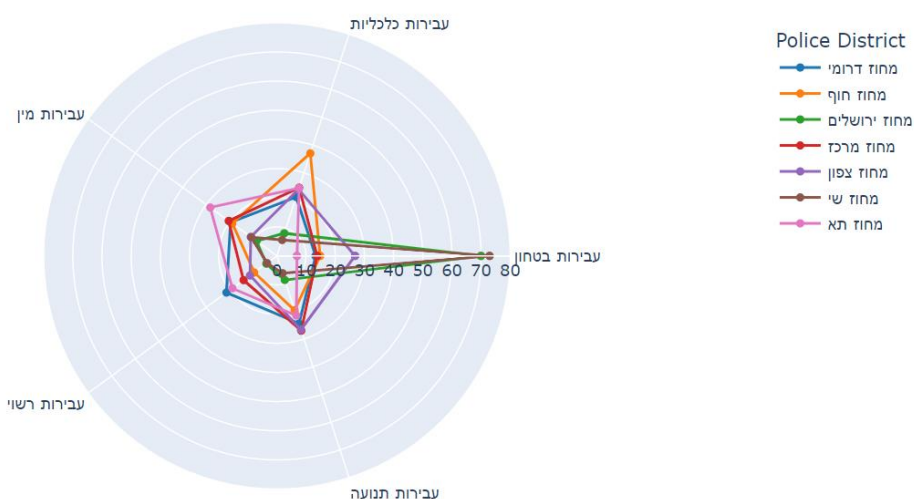
ביצענו אגרגציה לנתונים הגולמיים לפי רבעונים ו-Crime-Type, לאחר שסיננו אותם לעבירות כלכליות בלבד. הציר האופקי מתאר את הרבעונים באופן כרונולוגי, והאנכי מתאר את מספר התיקים, כאשר כל קבוצה מיוצגת על ידי צבע שונה (ומוצגת בגרף כ"עובי"). כשהקבוצות מוערמות זו על זו. בתרשים דאגנו שהשטחים הגדולים ביותר יוצגו בתחתית, כדי לאפשר ניתוח נוח וקל. עם זאת, החלטה זו מקשה על זיהוי מגמות בקבוצות הקטנות יותר, בעיה שנפתרת גם היא על ידי התרשים האינטראקטיבי. למשל - הסתרה של "עבירות פיסקליות" חושפת את העונתיות בעבירות על פיקוח כלכלי, שנראה שגדלות לרוב ברבעון הראשון:

Rise in Financial Crimes Is Primarily Driven by Fiscal Offenses



תרשים 4: החלק היחסי (באחוזים) של Crime-Groups שונים לפי מחוזות

Security Cases Predominate in Jerusalem and West Bank Districts



מה:

נתונים שחישבנו מתוך הנתונים הגולמיים. החלק היחסי (באחוזים) של מספר Crime-Groups, בכל המחוזות השונים. משתנה מספרי (האחוז היחסי של קבוצת העבירות במחוז) מתוך 5 קבוצות עבירות נבחרות (קטגוריאליות), בכל מחוז (קטגוריאליות).

למה:

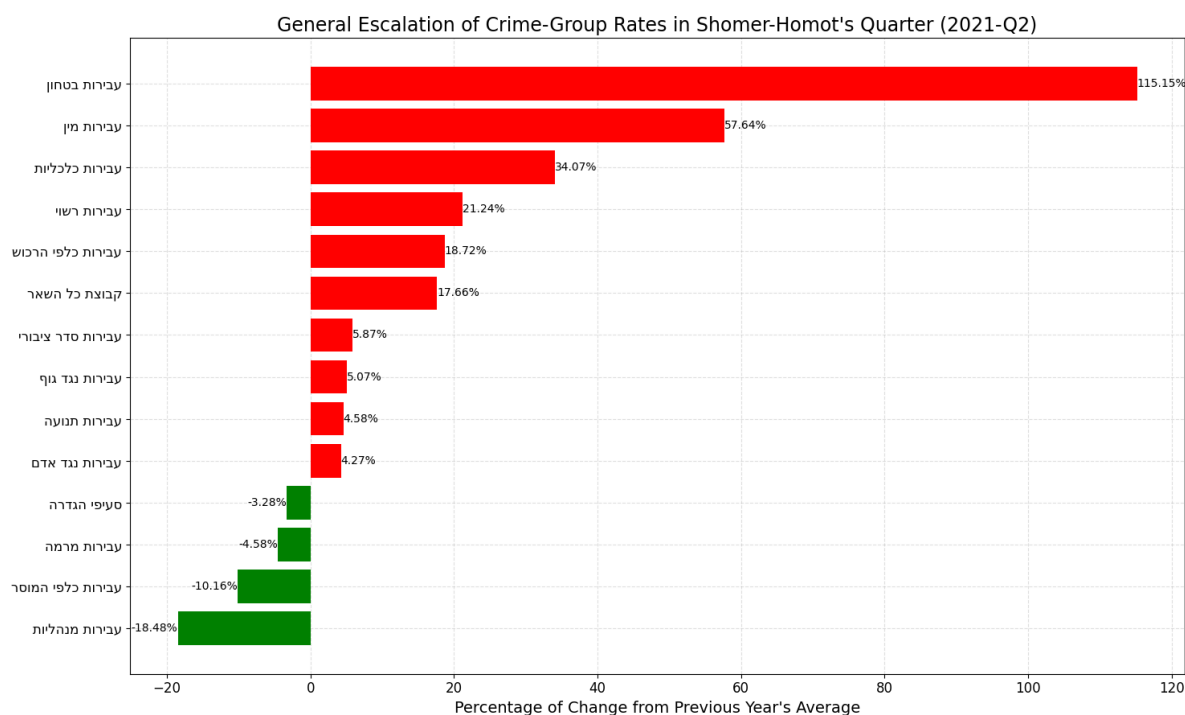
אנו מעוניינים להשוות בין המחוזות ולאפיין אותם אחד ביחס לשני - וגם בתוך עצמם. ההבנה אילו קבוצות-סוגי פשיעה נפוצות יותר באופן יחסי בכל מחוז תאפשר התאמה של אופי המחוז וכוח האדם, ציוד שנדרש לו ותשתיות חיוניות, ותקצוב מתאים.

איך:

גרף עכביש / רדאר אינטראקטיבי.

הגרף עושה שימוש בזוויות וברדיוס, שמאפשרים הצגה של משתנה קטגוריאליות ומספרי בהתאם. אנחנו הוספנו הפרדה של צבע שמאפיין כל פוליגון (מחוז), כך שמהתבוננות קצרה בצורות וצבעי הפוליגונים ניתן להבין אילו מחוזות מתמודדים עם אילו סוגי פשיעה יותר או פחות מאחרים. הגרף האינטראקטיבי מאפשר הסתרה והוספה של מחוזות כך שניתן גם לפשט את התרשים ולמנוע הסתרה של פוליגונים זה על ידי זה.

תרשים 5: אחוז השינוי בקבוצות-סוגי הפשיעה השונים ברבעון של שומר החומות



מה:

נתונים שחישבנו מתוך הנתונים הגולמיים, המייצגים את אחוז השינוי במספר התיקים שנפתחו בכל Crime-Group ברבעון של מבצע שומר החומות (2Q-2021) לעומת הממוצע בשנה הקודמת לו (2Q-2020) עד 1Q-2021). משתנה מספרי (אחוז השינוי ברבעון 2Q-2021) המחולק על פני 13 קבוצות-סוגי עבירות (קטגוריאלי).

למה:

עקב העלייה החדה שנצפתה ברבעון של שומר החומות בתרשים הראשון שהצגנו, אנו מעוניינים להציג את אחוז השינוי בקבוצות-סוגי הפשיעה השונות ברבעון של שומר החומות כדי לגלות האם הייתה מגמה כללית של עלייה/ירידה באחוזי הפשיעה לעומת השנה שקדמה למבצע, וגם להשוות את השינויים בכמות הפשיעה בין קבוצות-סוגי הפשיעה השונות.

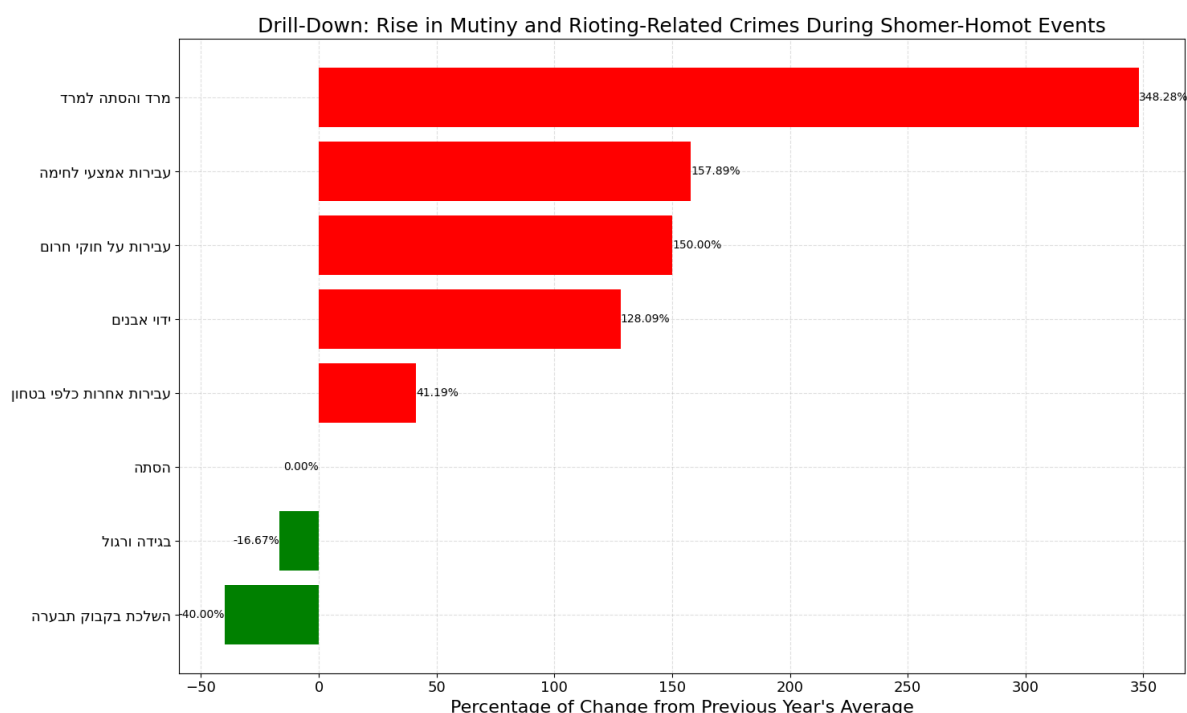
איך:

גרף עמודות מאוזן.

גרף זה משווה בין כלל קבוצות הפשיעה ומציג את אחוז השינוי ביחס לממוצע העבירות בשנה שקדמה למבצע (נקודת האפס בציר האופקי).

הצבעים בגרף עוזרים להבדיל בין עלייה לירידה בכמות הפשיעה- עלייה ברמת הפשיעה מיוצגת בצבע אדום (בהיותה אפקט שלילי) וירידה בצבע ירוק.

תרשים 6: אחוז השינוי בסוגי הפשיעה תחת עברות ביטחוניות ברבעון של שומר החומות



מה:

נתונים שחישבנו מתוך הנתונים הגולמיים, המייצגים את אחוז השינוי במספר התיקים שנפתחו בכל סוג פשיעה תחת ה-Crime-Group של עבירות ביטחוניות ברבעון של מבצע שומר החומות (2Q-2021) לעומת הממוצע בשנה הקודמת לו (2Q-2020 עד 1Q-2021). משתנה מספרי (אחוז השינוי ברבעון 2Q-2021) המחולק על פני 8 סוגי עבירות (קטגוריאלי).

למה:

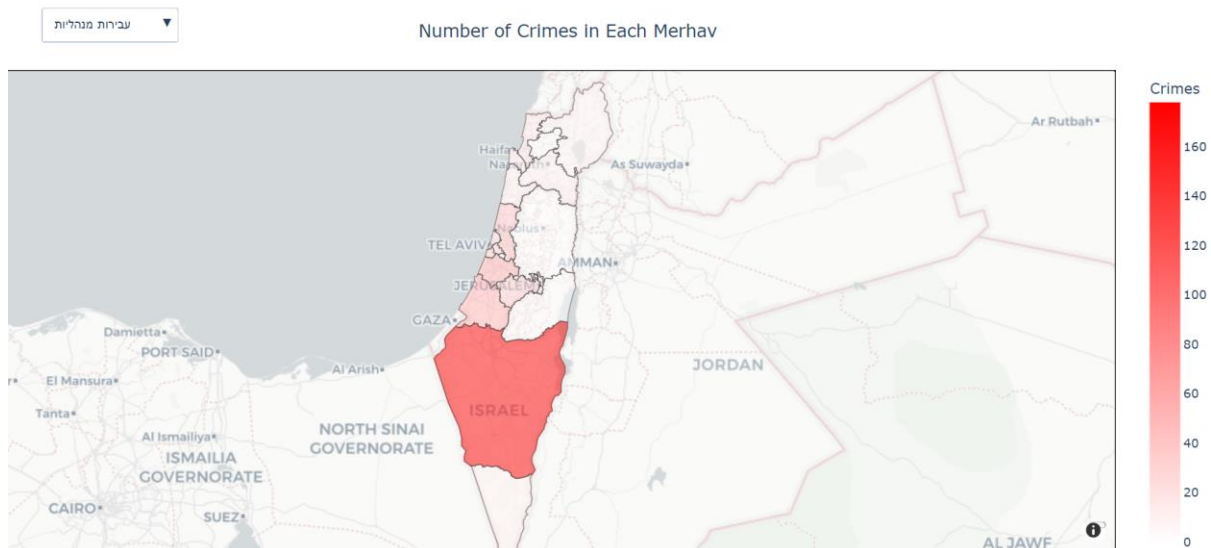
עקב העלייה החדה שנצפתה בעבירות הבטחוניות בתרשים הקודם שהצגנו, אנו מעוניינים להציג את אחוז השינוי בסוגי הפשיעה השונים תחת קבוצת העבירות הביטחוניות. המטרה היא להשוות בין השינויים בסוגים השונים ולגלות את מידת השינוי שלהם - האם הייתה עלייה בכל סוג עבירה? האם העלייה הייתה אחידה או שמא משמעותית יותר בסוגי פשיעה מסוימים?

איך:

גרף עמודות מאוזן.

גרף זה משווה בין כלל סוגי הפשיעה תחת עבירות בטחוניות ומציג את אחוז השינוי ביחס לממוצע העבירות בשנה שקדמה למבצע (נקודת האפס בציר האופקי). הצבעים בגרף עוזרים להבדיל בין עלייה לירידה בכמות הפשיעה - עלייה ברמת הפשיעה מיוצגת בצבע אדום וירידה בצבע ירוק.

תרשים 7: מפת מספר התיקים שנפתחו בכל מרחב



מה:

נתונים שחושבו מתוך הנתונים הגולמיים, והנתונים גיאוגרפיים של אזורי השיטור השונים. פוליגונים המתארים את שטחי האחריות של המרחבים על גבי מפה (נתונים גיאוגרפיים), מספר התיקים שנפתחו (מספרי) בכל מרחב (קטגוריאלי) מכל Crime-Group (קטגוריאלי, בתוספת קטגוריית סה"כ, סכימה של כל הקטגוריות).

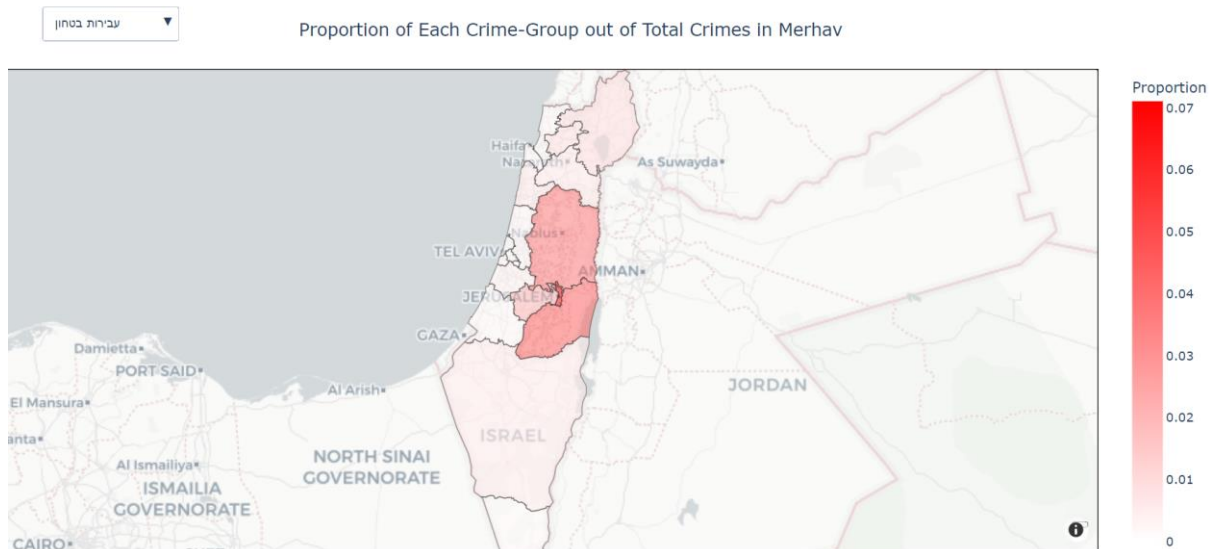
למה:

על מנת להבין מה העומס שחוזה כל מרחב משטרי, בסך הכל או בחתך של קבוצת-סוגי עבירות מסוימת, לצד הבנה של השטח שהוא חולש עליו והבנה של פריסת המרחבים בשטח. שימוש בתרשים יוכל לאפשר חלוקת משאבים והקצאת כוח אדם בהתאם למתואר במפה. (למשל, כמו שמתואר בצילום שלמעלה, מחוז נגב מתעסק עם עבירות מנהליות רבות וחולש על שטח רב, ייתכן שנדרשת הקצאת משאבים מתאימה).

איך:

מפה גיאוגרפית אינטראקטיבית עם סקאלת צבעים ו drop-down menu. הגרף עושה שימוש בצורות ובמיקום על מנת לתאר את הגיאוגרפיה של המרחבים, ובצבע על מנת לתאר את מספר התיקים במרחב. הצבע האדום מתאר מספר תיקים גדול, ומספר תיקים קטן שואף ללבן, שהוא ניטרלי לצבע המפה שתחתיו. אזור בו אין פשיעה - יראה כחלק במפה שנקי מהתערבות. מעבר לשימוש בתפריט הבחירה של Crime-Group, ניתן גם להעביר את העכבר מעל המרחבים השונים ולראות את שמותיהם ומספר התיקים המדויק שנפתחו בהם.

תרשים 8: מפת פרופורציית Crime-Group בכל מרחב



מה:

נתונים שחושבו מתוך הנתונים הגולמיים, והנתונים גיאוגרפיים של אזורי השיטור השונים. פוליגונים המתארים את שטחי האחריות של המרחבים על גבי מפה (נתונים גיאוגרפיים), פרופורציית קבוצת-סוגי הפשיעה (מספרי) בכל מרחב (קטגוריאל) מכל Crime-Group (קטגוריאל).

למה:

חסרון של המפה הקודמת הוא שמרחבים גדולים (בשטח, אוכלוסייה, כוחות שיטור) יקבלו לרוב ערכים גדולים יותר במספרי התיקים. התייחסות זו עשויה להציג אזורים "קטנים" כנטולי פשיעה ותקשה על הבנה של אופי הפשיעה בהם. על כן, החלטנו להציג מפה נוספת בה מספר התיקים מקבוצת-סוג פשיעה מסוים מוצג כפרופורציה מתוך כלל התיקים במרחב (בדומה לפרופורציה המתוארת בתרשים 4). בעזרת המפה ניתן לראות איפה בארץ יותר או פחות נפוצים סוגי פשיעה שונים באופן יחסי. למשל, בדוגמה למעלה, החלק היחסי של עבירות ביטחוניות יותר משמעותי במרחבים שומרון, יהודה, וקדם.

איך:

מפה גיאוגרפית אינטראקטיבית עם סקאלת צבעים ו-drop-down menu. הגרף עושה שימוש בצורות ובמיקום על מנת לתאר את הגיאוגרפיה של המרחבים, ובצבע על מנת לתאר את מספר התיקים היחסי. הצבע האדום מתאר מספר תיקים גדול, ומספר תיקים קטן שואף ללבן, שהוא ניטרלי לצבע המפה שתחתיו. אזור בו אין פשיעה מהסוג הנבחר - יראה כחלק במפה שנקי מהתערבות. מעבר לשימוש בתפריט הבחירה של Crime-Group, ניתן גם להעביר את העכבר מעל המרחבים השונים ולראות את שמותיהם ויחס התיקים מהסוג הנבחר שנפתחו בהם.

A pizza is basically a real-time pie chart of how much pizza is left.

