

מתקפות טרור בישראל לאורך השנים 1970-2017

ויזואליזציה של מידע - 382-1-3203

ויזואליזציה של מידע: 382.1.3203

מגישים: ירין שוחט, מקסים ליסיאנסקי, רועי קרמר ואברהם אלבז

[](https://www.kaggle.com/datasets/START-UMD/gtd/data)תמונה שמכילה שחור, חשיכה

התיאור נוצר באופן אוטומטי

Data

GitHub

Dashboard

תאריך הגשה: XXXXX

**מבוא**

בחרנו לעבד, לנתח ולהציג נתונים אודות אירועי טרור בישראל לאורך השנים, מתוך מאגר הנתונים Global Terrorism Database (GTD).

אנו נשתמש במאגר נתונים רחב ומקיף המכיל מידע על אירועי טרור בין השנים 1970 ל-2017.

הנתונים נלקחו מאתר Kaggle ומתוכם 2,183 שורות המתארות אירועי טרור שהתרחשו בשטח ישראל, לצד 135 עמודות המכילות פרטים מגוונים על כל אירוע, כולל תאריך האירוע, מיקום, מספר המחבלים, מספר הנפגעים, סוג התקיפה, המטרה ועוד. חשוב לציין שלא נעשה שימוש בכל 135 העמודות, אלא רק בעמודות הרלוונטיות לשאלת המחקר, על מנת להתמקד בניתוח נתונים משמעותיים.

בחרנו להתמקד בפיגועים שהתרחשו בשטח ישראל בלבד, במטרה להבין את דפוסי הפעולה והמאפיינים הייחודיים למרחב המקומי.

**השאלה המרכזית עליה נרצה לענות בעזרת הנתונים שבחרנו היא:**

כיצד השתנה דפוס הטרור בישראל לאורך השנים מבחינת תדירות האירועים, הפיזור הגיאוגרפי ושיטות הפעולה?

מטרתנו בפרויקט היא להבין את מגמות הטרור בישראל לאורך זמן, תוך התמקדות בזיהוי מוקדי הפעולה המרכזיים, קשרים בין מאפייני האירועים והתפתחותם ההיסטורית.

יתר על כן, נרצה לענות על מספר שאלות משנה נוספות:

1. תדירות התקיפות לפי אזורים: מהם האזורים בהם התרחשו המספר הגבוה ביותר של תקיפות טרור?
2. קורלציה בין מספר המחבלים למספר הנפגעים: האם קיים קשר בין כמות המחבלים באירוע לכמות הנפגעים?
3. כמות פיגועי הטרור לאורך השנים: כיצד השתנתה תדירות הפיגועים לאורך השנים?

**נתונים**

מאגר המידע שלנו מכיל 2,183 רשומות ו-19 עמודות, אשר כל רשומה מייצגת תצפית של אירוע טרור בישראל, עמודת 'eventid' מהווה מפתח מזהה בין כל אירוע וכל עמודה נוספת היא פיצ'ר אודות האירוע עצמו אשר מביאה לנו מידע אודות האירוע כגון מספר הנפגעים, מיקום, סוג הפיגוע ועוד.

Dataset Type: Flat Table

Item: תצפית של אירוע טרור

|  |  |  |  |
| --- | --- | --- | --- |
| **Cardinality** | **Meaning** | **Type** | **Attribute** |
| 2183 Possible values | 8 הספרות הראשונות מייצגות תאריך, 4 הספרות האחרונות מייצגות ID ייחודי | Categorical (nominal) | eventid  (Primary Key) |
| 46 Possible values שנים: 1971-2017 | השנה שבה התרחש האירוע | Ordered -Quantitative (Sequential) | iyear |
| 12 Possible values 12 חודשי השנה | החודש שבו התרחש האירוע | Ordered -Quantitative (Cyclic) | imonth |
| 32 Possible values 31 ימי החודש האפשריים + 0 | היום שבו התרחש האירוע | Ordered -Quantitative (Cyclic) | iday |
| 1 Possible value Isreal | קוד מדינה | Categorical (nominal) | country |
| 369 Possible values | שם העיר, הכפר או העיירה שבה התרחש האירוע | Categorical (nominal) | city |
| 418 Possible values | קו אורך | Ordered -Quantitative (Diverging) | latitude |
| 422 Possible values | קו רוחב | Ordered -Quantitative (Diverging) | longitude |
| 14 Possible values  0-30 | המספר הכולל של המחבלים המשתתפים באירוע | Ordered -Quantitative (Sequential) | nperps |
| 30 Possible values 0-42 | מספר ההרוגים הכולל באירוע | Ordered -Quantitative (Sequential) | nkill |
| 70 Possible values 0-192 | כמות פצועים | Ordered -Quantitative (Sequential) | nwound |
| 232 Possible values | פירוט המדינה והעיר בה התבצעה ההתקפה | Categorical (nominal) | location |
| 2 Possible values 0-לא -כן1 | האם ההתקפה הצליחה | Categorical (nominal) | success |
| 9 Possible values | הדרך או הנשק בעזרתו נעשתה ההתקפה | Categorical (nominal) | attacktype1 |
| 2 Possible values 0-לא -כן1 | האם ההתקפה התבצעה באמצעות התאבדות | Categorical (nominal) | suicide |
| 20 Possible values | סוג המטרה/קורבן | Categorical (nominal) | targtype1 |
| 8 Possible values | סוג הנשק בשימוש באירוע | Categorical (nominal) | weaptype1\_txt |
| 81 Possible values | שם הקבוצה שביצעה את הפיגוע | Categorical (nominal) | gname |
| 2 Possible values 0-לא -כן1 | האם ההתקפה נמשכה פחות מ24 שעות או יותר | Categorical (nominal) | extended |

* לבדוק לעומק את הATTRIBUTE TYPES
* להוסיף בקרדינליות את האופציות

**מטלות**

1. **תיאור מטלות המשתמש במונחי התחום:**
2. התפלגות תקיפות הטרור לפי אזורים: כמות אירועי הטרור לפי מיקומים שונים בישראל?  
   התפלגות תקיפות טרור לפי אזורים בישראל מתארת מטלה בתחום ביטחון לאומי וניתוח אירועים גיאוגרפים. במטלה זאת נרצה להשוות בין כמות מתקפות הטרור באזורים שונים בישראל ולהבחין במקומות בארץ בהם יש יותר פיגועים, בכך נוכל לזהות מיקומים הדורשים צורך בהיערכות ביטחונית מוגברת.
3. קורלציה בין מספר המחבלים למספר הנפגעים: האם קיים קשר בין כמות המחבלים באירוע לכמות הנפגעים?

המטרה המרכזית של ניתוח הקורלציה היא לזהות קשרים בין המשתנים: מספר המחבלים (nperps), מספר ההרוגים (nkill), ומספר הפצועים (nwound). ניתוח זה מאפשר להבין האם קיים קשר ליניארי בין היקף האירוע (מספר המחבלים) לבין התוצאה האנושית (הרוגים ופצועים), ולהעריך האם אירועים חמורים יותר מתאפיינים במאפיינים מסוימים.

1. כמות פיגועי הטרור לאורך השנים: כיצד השתנתה תדירות הפיגועים לאורך השנים?

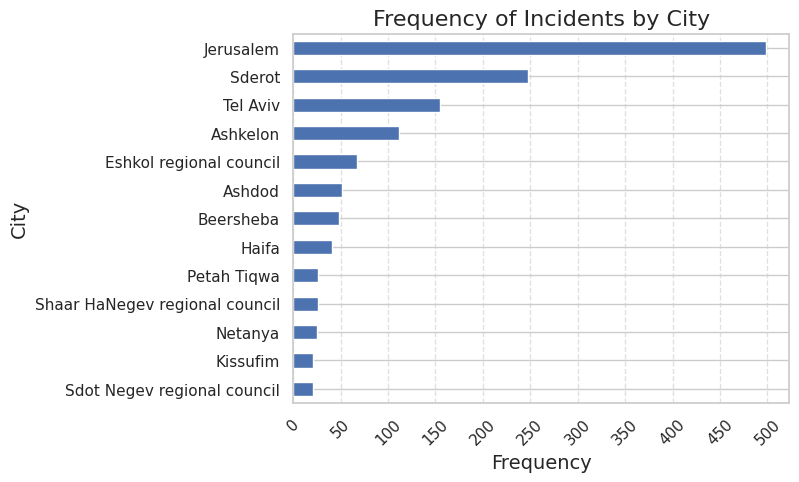
הבנת השינויים בתדירות פיגועי הטרור במדינת ישראל לאורך השנים במטרה לזהות האם ישנה קורלציה או מגמה עם השנים.

|  |  |  |
| --- | --- | --- |
| Target | Action | Real world task |
| Distribution | Search - Explore | התפלגות תקיפות הטרור לפי אזורים |
| Correlation | Compare | קורלציה בין מספר המחבלים למספר הנפגעים: האם קיים קשר בין כמות המחבלים באירוע לכמות הנפגעים? |
| Trends | Compare | ניתוח כמות פיגועי הטרור לאורך השנים והבנת מגמות שינוי בתדירות הפיגועים |

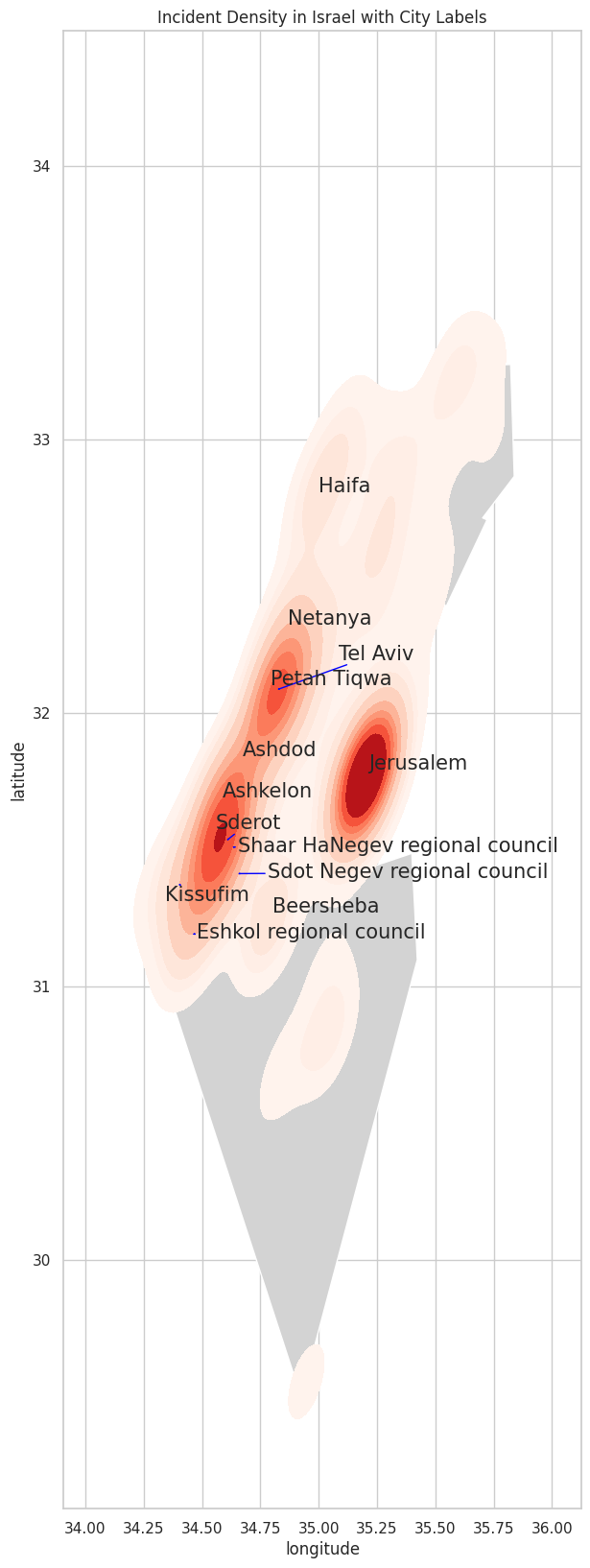
1. **מיפוי המטלות במונחים האבסטרקטיים בהתאם לטיפולוגיה של Munzner:**

**עיצובים חלופיים**

1. התפלגות תקיפות הטרור לפי אזורים: כמות אירועי הטרור לפי מיקומים שונים בישראל?  
   התפלגות תקיפות טרור לפי אזורים בישראל מתארת מטלה בתחום ביטחון לאומי וניתוח אירועים גיאוגרפים. במטלה זאת נרצה להשוות בין כמות מתקפות הטרור באזורים שונים בישראל ולהבחין במקומות בארץ בהם יש יותר פיגועים, בכך נוכל לזהות מיקומים הדורשים צורך בהיערכות ביטחונית מוגברת.

**חלופה 1 למטלה 1:**

* **Marks:**  
  קווים – כל פס מייצג עיר
* **Channels:**  
  מיקום בציר האופקי מנקודת התחלה שווה  
  אורך הקו – מספר הפיגועים שהתרחשו באותו העיר  
  מיקום בציר האנכי בסדר ממוין לפי כמות הפיגועים בסדר עולה
* **אקספרסיביות:**  
  השימוש בגרף עמודות מתאים לסוג הנתונים, תכונה קטגוריאלית(עיר) ותכונה כמותית(תדירות התקיפות)  
  הערוץ הוויזואלי של אורך העמודה מתאים ומייצג את הכמות
* **אפקטיביות:**  
  קל להשוות בין התדירות של הפיגועים בין הערים השונות  
  המיקום של העיר בציר ה-y עוזר לעין לזהות את הגדלים של תדירות הפיגועים  
  קווים מקווקוים שיוצאים מציר ה-x שעוזרים להעריך את הכמות של כל עיר

**חלופה 2 למטלה 1:**

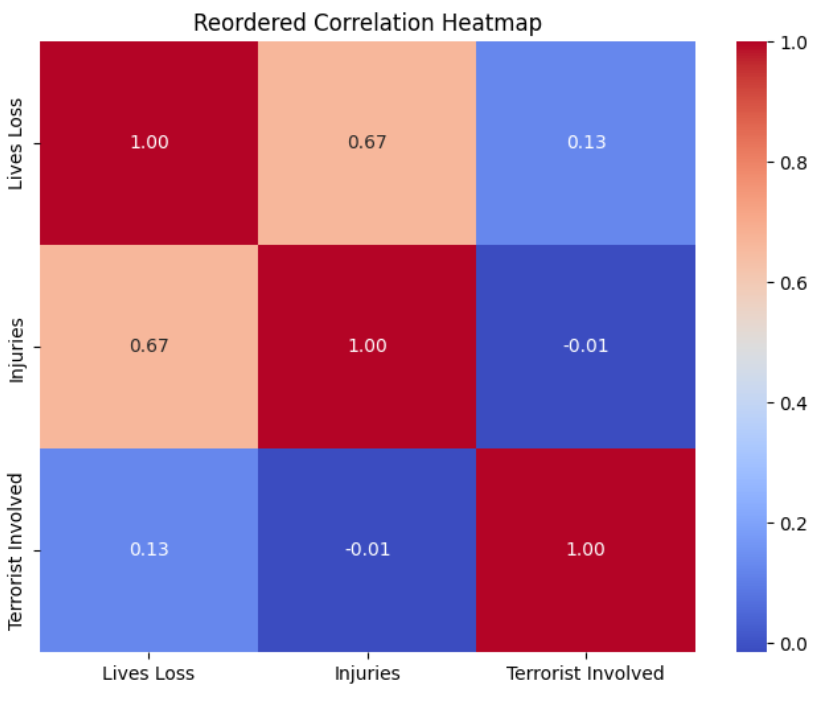
* **Marks:**  
  נקודות – כל נקודה מייצגת פיגוע
* **Channels:**  
  Spatial Region – מיקום בציר האופקי והאנכי לפי נ"צ  
  Color Saturation – הצבע מייצג את כמות הפיגועים שהיו באותו אזור, ככל שכהה יותר ככה יש יותר פיגועים  
  Area – גודל האליפסה מסמל את כמות הפיגועים באותו אזור
* **אקספרסיביות:**הגרף מתאים לנתונים, שני תכונות כמותיות(קו אורך ורוחב) שמייצגות את הצירים ותכונה כמותית(תדירות התקיפות) שמייצגת את הצבע  
  הערוץ הוויזואלי של הצבע מתאים והוא מייצג את תדירות התקיפות לפי הרוויה של הצבע
* **אפקטיביות:**מפת החום עוזרת להבחין במקומות בהם יש ריכוז גבוה יותר של מקומות  
  בחירת הצבעים בתור הצבע החזק לאזור שבו יש הרבה פיגועים מושכת את העין לעומת האזורים עם הצבעים החלשים שבהם אין הרבה פיגועים  
  תוויות הערים עוזר לזהות בערך את המיקום של הערים

|  |  |  |
| --- | --- | --- |
|  | חלופה 1 – גרף עמודות היסטוגרמה | חלופה 2 – גרף מפה של צפיפות |
| יתרונות | מהיר לקריאה, אפשר ישירות לדעת באיזו עיר יש הכי הרבה פיגועים. אפשר להעריך את הכמות של הפיגועים בכל עיר. | אפשר לזהות אזורים כללים בארץ שבהם יש הרבה פיגועים. זיהוי מהיר בעזרת הצבע של מקומות שבהם יש הרבה פיגועים. |
| חסרונות | מוגבל רק לערים שמופיעות שם ולא לשאר המקומות בארץ. יש הגבלה של מספר הערים שאפשר להציג, יותר מידי ערים היה מעמיס על הגרף. | קשה לדעת את הכמות המדויקת של הפיגועים בכל מקום. |

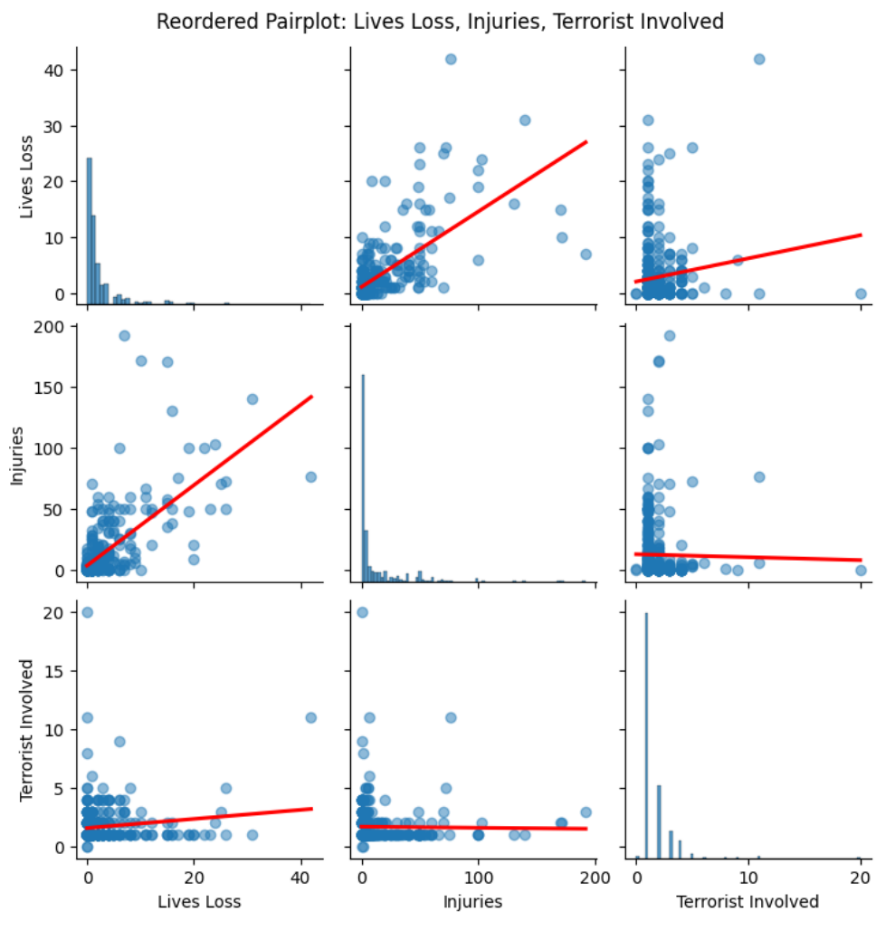
**הוויזואליזציה שנבחרה היא:** X

1. קורלציה בין מספר המחבלים למספר הנפגעים: האם קיים קשר בין כמות המחבלים באירוע לכמות הנפגעים?

המטרה המרכזית של ניתוח הקורלציה היא לזהות קשרים בין המשתנים: מספר המחבלים (nperps), מספר ההרוגים (nkill), ומספר הפצועים (nwound). ניתוח זה מאפשר להבין האם קיים קשר ליניארי בין היקף האירוע (מספר המחבלים) לבין התוצאה האנושית (הרוגים ופצועים), ולהעריך האם אירועים חמורים יותר מתאפיינים במאפיינים מסוימים.

**חלופה 1 למטלה 2:**

* **Marks:**  
  כל תא במפת החום הוא סימון (Mark) שמציג את הקורלציה בין שני משתנים: מספר מחבלים, מספר הרוגים, ומספר פצועים
* **Channels:**  
  צבע התא - מייצג את ערך הקורלציה (מסקלת 1- עד 1). צבעים חמים (אדום) מייצגים קורלציה חיובית גבוהה, וצבעים קרים (כחול) מייצגים קורלציה שלילית  
  ציר X ו-Y - מייצגים את המשתנים עצמם (nperps, nkill, nwound)  
  מספרים על התאים - ערך הקורלציה המספרי המדויק בין המשתנים
* **אקספרסיביות:**  
  מציגה עוצמת קשר ליניארי בלבד: הערכים והצבעים מתארים את עוצמת הקורלציה בלבד, מבלי לפרט על צורת הקשר, על חריגות או על פיזור התצפיות. זה מגביל את המידע שאפשר להפיק מהמפה.
* **אפקטיביות:**מתאימה להצגה מהירה של עוצמת הקשרים הליניאריים: מפת החום נותנת תמונה ברורה ומהירה על חוזק ועוצמת הקשרים בין משתנים שונים באמצעות צבעים וערכים מספריים. זה עוזר לנתח מערכים גדולים בלי להתעמק בפרטי כל קשר.

**חלופה 2 למטלה 2:**

* **Marks:**  
  נקודות (Points) - כל נקודה בגרף מייצגת זוג ערכים מתצפית אחת (למשל, מספר מחבלים מול מספר הרוגים או מספר פצועים). הנקודות הן הסימנים העיקריים בגרף.  
  קווים אדומים (Red Lines) - הקווים האדומים מייצגים את המגמה הכללית בין שני המשתנים באמצעות קו רגרסיה ליניארי, שמראה את כיוון הקשר (חיובי, שלילי או אפסי).  
  מה עם ה-BarPlot?
* **Channels:**  
  מיקום הנקודות (Position):   
   ציר X: מציג את הערכים של המשתנה הראשון בכל זוג (לדוגמה, מספר המחבלים - nperps).  
   ציר Y: מציג את הערכים של המשתנה השני בכל זוג  
  צפיפות הנקודות (Density):   
  מקומות עם ריכוז גבוה של נקודות מצביעים על תצפיות שחוזרות על עצמן בטווחים מסוימים.  
  שיפוע הקו האדום (Slope of the Red Line):  
  קו בעל שיפוע חיובי מצביע על קשר חיובי (כשערך משתנה אחד עולה, גם השני עולה), קו בעל שיפוע אפסי או כמעט אפסי מעיד על קשר חלש או היעדר קשר.  
  סקלה בצירים (Axis Scales):  
  הסקלה בכל ציר מספקת את טווח הערכים של המשתנים, מה שמאפשר להבין את פיזור הנתונים.
* **אקספרסיביות:**  
  מציגה קשרים מורכבים, פיזור וחריגות: באמצעות פיזור הנקודות ניתן לראות אם הקשר חזק או חלש, להבין אילו ערכים חוזרים על עצמם ואילו משתנים מתנהגים כחריגים. היא גם מאפשרת זיהוי קשרים מורכבים שאינם בהכרח ליניאריים.
* **אפקטיביות:**מתאימה לניתוח מעמיק של קשרים ומגמות: מטריצת גרף פיזור מאפשרת להבין לא רק האם קיים קשר בין שני משתנים, אלא גם את אופי הקשר (ליניארי, לא ליניארי) ואת פיזור הנתונים. בעזרת קווי המגמה, ניתן לזהות את כיוון הקשר ולזהות אם קיימות קבוצות חריגות שדורשות בחינה נוספת.

|  |  |  |
| --- | --- | --- |
|  | חלופה 1 – מפת חום של קורלציה | חלופה 2 – מטריצת גרף פיזור |
| יתרונות | קלה לקריאה ומהירה להבנה: באמצעות שימוש בצבעים ובערכים מספריים, ניתן להבין בקלות אילו משתנים קשורים אחד לשני בעוצמה חזקה או חלשה. זה מתאים במיוחד למערכי נתונים גדולים ורחבים. | מתאימה לזיהוי דפוסים מורכבים ונקודות חריגות: הגרף עוזר לאתר תצפיות חריגות ולבחון האם ישנם קשרים לא ליניאריים בין המשתנים. יתרון נוסף הוא היכולת להתמקד בזוג משתנים מסוים ולנתח אותו לעומק בעזרת הפיזור והקו המנחה. |
| חסרונות | מוגבלת רק לקשרים ליניאריים: אינה מתארת קשרים מורכבים או מגמות לא ליניאריות, וגם אינה חושפת מידע על חריגות או פיזור בתוך הקשרים. היא מתאימה בעיקר לשימוש ראשוני לבדיקת קשרים כלליים. | עשויה להיות עמוסה עם נתונים רבים: כאשר יש כמות גדולה של משתנים או כאשר הנתונים צפופים, קשה מאוד לקרוא את מטריצת הפיזור. יש גם צורך בניתוח פרטני של כל זוג משתנים, מה שדורש זמן רב. |

**הוויזואליזציה שנבחרה היא:** X

1. כמות פיגועי הטרור לאורך השנים: כיצד השתנתה תדירות הפיגועים לאורך השנים?

הבנת השינויים בתדירות פיגועי הטרור במדינת ישראל לאורך השנים במטרה לזהות האם ישנה קורלציה או מגמה עם השנים.

**חלופה 1 למטלה 3:**



* **Marks:**  
  **שטחים (Areas)**: מפת החום משתמשת במלבנים המייצגים את החיתוך בין "שנה" ( ציר ה (X ו-"מספר תקריות) "ציר ה(Y
* **Channels:**  
  1. מיקום (Position):
  + אופקי: מייצג את השנה על ציר ה-X.
  + אנכי: מייצג את מספר התקריות (למרות שבמפה זו יש שורה אחת בלבד כי מדובר בסיכום שנתי).

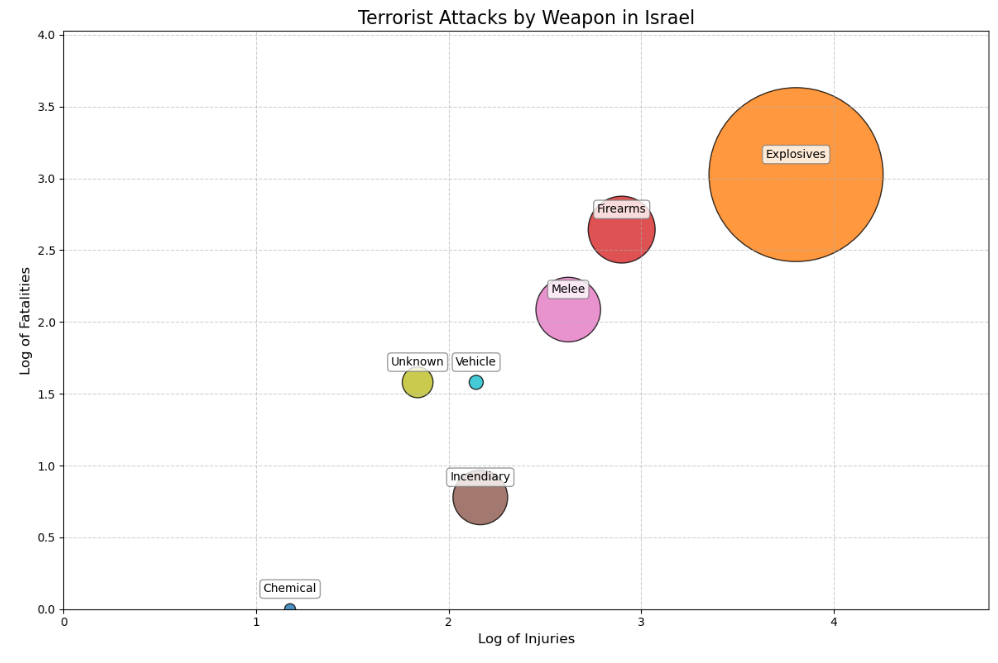
2. צבע (Color):

* + עוצמת הצבע בכל מלבן מקודדת את מספר התקריות.
  + צבע כהה מציין מספר גבוה יותר של תקריות, וצבע בהיר מציין מספר נמוך יותר.

3. גודל (Size) (מובנה ברשת):

* + גודל המלבנים אחיד ומבטיח פרשנות עקבית בין השנים, אך מבנה הרשת מארגן את המידע לפי שנים.
* **אקספרסיביות:**  
  מפת החום משתמשת בעקרון ההבעה על ידי התאמת ערוצי הוויזואליזציה לסוגי הנתונים. נתונים מסודרים, כמו שנים, מוצגים בציר ה-X בסדר רציף ולוגי. עוצמת מספר התקריות מקודדת באמצעות עוצמת הצבע, ערוץ כמותי, שבו צבע כהה מציין ערכים גבוהים וצבע בהיר מציין ערכים נמוכים. כך הוויזואליזציה אינטואיטיבית ומעבירה דפוסים בנתונים בצורה מדויקת, בהתאם לעיקרון שעבור כל סוג נתונים יש לבחור בערוץ מתאים.
* **אפקטיביות:**המיקום בציר האופקי (שנה) מנצל ערוץ מדויק ובולט עבור נתונים מסודרים, מה שמאפשר למשתמשים לעקוב בקלות אחר שינויים לאורך זמן. בנוסף, עוצמת הצבע מקודדת בצורה יעילה את עוצמת התקריות, כאשר צבעים כהים מייצגים באופן אינטואיטיבי ערכים גבוהים. בחירות אלו מקלות על זיהוי דפוסים, זיהוי נקודות שיא והשוואה בין שנים, וממקסמות דיוק, יכולת הבחנה וקיבוץ עבור המערכת הקוגניטיבית האנושית.

**חלופה 2 למטלה 3:**



* **Marks:**  
  **נקודות**: כל בועה מייצגת קטגוריה של סוג הנשק שבו נעשה שימוש בפיגועי טרור.
* **Channels:**

המיקום בציר ה-X מייצג את לוג הפצועים, בעוד המיקום בציר ה-Y מייצג את לוג ההרוגים, כך שהגרף מאפשר לבחון את הקשרים בין כמות הפצועים לכמות ההרוגים עבור כל קטגוריה. גודל הבועה משקף את מספר הפיגועים בכל קטגוריה של נשק, ומדגיש את ההבדלים בין הקטגוריות מבחינת היקף ההשפעה שלהן. בנוסף, כל קטגוריה מיוצגת בצבע ייחודי, המקל על ההבחנה בין הקטגוריות ומסייע בזיהוי מהיר וברור של כל סוג נשק.

* **אקספרסיביות:**

הגרף מתאים לנתונים בכך שהוא משתמש בשתי תכונות כמותיות, לוג של פצועים ולוג של הרוגים, כצירים, ותכונה קטגורית, סוג נשק, כדי להבדיל בין הבועות. שטח הבועה מייצג באופן ברור את מספר הפיגועים, מה שמאפשר השוואה בין הגדלים של הקטגוריות השונות. בנוסף, הערוצים הוויזואליים שבגרף מחזקים את הבהירות שלו: השימוש בצבעים ייחודיים לכל קטגוריה מבטיח הבחנה ברורה בין הקטגוריות, והשימוש בסקאלה לוגריתמית בשני הצירים מדגיש את הפערים בין הנתונים, במיוחד עבור קטגוריות עם ערכים שונים משמעותית.

* **אפקטיביות:**

הגרף מספק תובנות והשוואות ברורות ומאפשר לזהות בקלות קטגוריות עם מספר גבוה של הרוגים ופצועים, כגון קטגוריית "Explosives". בועות קטנות יותר, הממוקמות נמוך בצירים, מייצגות קטגוריות עם פחות פיגועים, הרוגים ופצועים. השימוש בסקאלה לוגריתמית בגרף הכרחית, שכן הדבר מדגיש את ההבדלים בין קטגוריות של נשקים. תוויות הממוקמות מעל הבועות מקלות על זיהוי הקטגוריות באופן ישיר, ללא צורך בLEGEND נפרד. בנוסף, השימוש בצבעים ברורים ומובחנים תורם למניעת בלבול ומקל על ההבחנה בין הקטגוריות השונות.

|  |  |  |
| --- | --- | --- |
|  | חלופה 1 – ש | חלופה 2 – ש |
| יתרונות | ש | ש |
| חסרונות | ש | ש |

**הוויזואליזציה שנבחרה היא:** X