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

Data

Dashboard

GitHub

תאריך הגשה: XXXXX

תוכן עניינים

מבוא 3
נתונים 4
מטלות 5
האם ישנם אזורים בישראל עם ריכוז יוצא דופן של מתקפות טרור בהשוואה לאחרים? 5
קורלציה בין מספר המחבלים למספר הנפגעים: האם קיים קשר בין כמות המחבלים באירוע לכמות הנפגעים? 5
מהם ההבדלים בין סוגי הנשק בתרומתם לחומרת האירועים(מספר נפגעים והרוגים) ותדירות האירועים מכל סוג נשק? 6
עיצובים חלופיים 7
האם ישנם אזורים בישראל עם ריכוז יוצא דופן של מתקפות טרור בהשוואה לאחרים? 7
קורלציה בין מספר המחבלים למספר הנפגעים: האם קיים קשר בין כמות המחבלים באירוע לכמות הנפגעים? 10
מהם ההבדלים בין סוגי הנשק בתרומתם לחומרת האירועים(מספר נפגעים והרוגים) ותדירות האירועים מכל סוג נשק? 14
הסבר על העיצובים שנבחרו ועל יישומם 18

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.

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. האם ישנם אזורים בישראל עם ריכוז יוצא דופן של מתקפות טרור בהשוואה לאחרים?
2. קורלציה בין מספר המחבלים למספר הנפגעים: האם קיים קשר בין כמות המחבלים באירוע לכמות הנפגעים?
3. מהם ההבדלים בין סוגי הנשק בתרומתם לחומרת האירועים(מספר נפגעים והרוגים) ותדירות האירועים מכל סוג נשק?

נתונים

מאגר המידע שלנו מכיל 2,183 רשומות, לא נעשה שימוש בכל 135 העמודות, אלא רק בעמודות הרלוונטיות לשאלת המחקר, לכן המאגר שלנו מכיל 11 עמודות. 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

Attribute	Type	Meaning	Cardinality
eventid (Primary Key)	Categorical (nominal)	8 הספרות הראשונות מייצגות תאריך, 4 הספרות האחרונות מייצגות ID ייחודי	2183 Possible values
iyear - Year	Ordered - Quantitative	השנה שבה התרחש האירוע	46 Possible values שנים: 1971-2017
imonth - Month	Ordered - Quantitative	החודש שבו התרחש האירוע	12 Possible values 12 חודשי השנה
iday - Day	Ordered - Quantitative	היום שבו התרחש האירוע	31 Possible values 31 ימי החודש האפשריים
city	Categorical (nominal)	שם העיר, הכפר או העיירה שבה התרחש האירוע	369 Possible values
latitude	Ordered - Quantitative	קו אורך	418 Possible values
longitude	Ordered - Quantitative	קו רוחב	422 Possible values
nperps - Terrorists Number	Ordered - Quantitative	המספר הכולל של המחבלים המשתתפים באירוע	14 Possible values 0-30
nkill - Deaths	Ordered - Quantitative	מספר ההרוגים הכולל באירוע	30 Possible values 0-42
nwound - Injuries	Ordered - Quantitative	כמות פצועים	70 Possible values 0-192
weaptype1_txt - Weapon Type	Categorical (nominal)	סוג הנשק בשימוש באירוע	8 Possible values

מטלות

1. **האם ישנם אזורים בישראל עם ריכוז יוצא דופן של מתקפות טרור בהשוואה לאחרים?**

א. מונחי התחום:

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.

ב. מונחי Munzner:

Target	Action	Real world task
Features	Analyze → Produce → Derive	קיבוץ הפיגועים לפי המיקום בו הם התרחשו
Distribution	Search → Browse	איך מתפלגת כמות התקיפות בישראל לפי מיקומים גיאוגרפים שונים?
Extremes	Search → Locate	האם ישנם אזורים בישראל עם ריכוז יוצא דופן של מתקפות טרור בהשוואה לאחרים?
Extremes	Query → Identify	זיהוי אזורים חריגים בישראל בהם יש ריכוז יוצא דופן של מתקפות טרור

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

א. מונחי התחום:

המטרה המרכזית של ניתוח הקורלציה היא לזהות קשרים בין המשתנים: מספר המחבלים (Terrorists), מספר ההרוגים (Deaths), ומספר הפצועים (Injuries). ניתוח זה מאפשר להבין האם קיים קשר ליניארי בין היקף האירוע (מספר המחבלים) לבין התוצאה האנושית (הרוגים ופצועים).

ב. מונחי Munzner:

Real world task	Action	Target
ניתוח הקשרים בין מספר המחבלים, מספר ההרוגים ומספר הפצועים כדי לזהות מגמות כלליות.	Analyze → Consume → Discover	Features
האם קיים קשר חזק בין מספר המחבלים לבין כמות הנפגעים (הרוגים ופצועים).	Search → Browse	Correlation
השוואת הקשרים בין משתנים שונים כדי לבדוק האם מספר המחבלים יכול לנבא את מספר הפצועים וההרוגים.	Query → Compare	Correlation

3. מהם ההבדלים בין סוגי הנשק בתרומתם לחומרת האירועים(מספר נפגעים והרוגים) ותדירות האירועים מכל סוג נשק?

א. מונחי התחום:

נרצה להשוות בין סוגי נשק שונים שבהם המחבלים השתמשו בכל פיגוע, בהתחשבות במספר הנפגעים וההרוגים וכמות הפיגועים שבוצעו בכל סוג נשק. נרצה גם להבחין בסוגי נשק קטלניים יותר(גורמי נזק).

ב. מונחי Munzner:

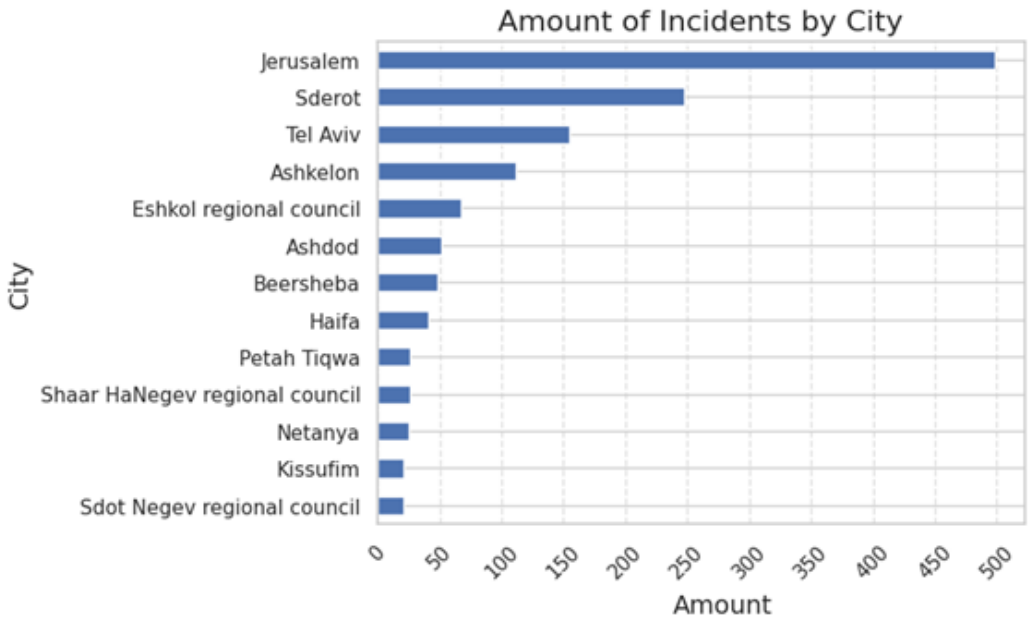
Real world task	Action	Target
ניתוח ההבדלים בין סוגי הנשק והשפעתם על מספר הנפגעים (הרוגים ופצועים), תוך זיהוי אילו סוגי נשק קטלניים יותר.	Analyze → Consume → Discover	Trends
השוואת קטגוריות שונות של כלי נשק מבחינת מספר הפצועים וההרוגים כדי להבין אילו נשקים גורמים לנפגעים רבים יותר.	Search → Browse	Correlation
בחינת הקשרים בין סוגי הנשק לבין מספר הנפגעים (פצועים והרוגים) כדי לזהות האם יש קשר ברור בין סוג הנשק להיקף הפגיעות.	Query → Compare	Distribution

עיצובים חלופיים

1. האם ישנם אזורים בישראל עם ריכוז יוצא דופן של מתקפות טרור בהשוואה לאחרים?

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 (Position on common scale)

אורך הקו – מספר הפיגועים שהתרחשו באותו העיר

מיקום בציר האנכי מייצג את הערים בסדר ממוין לפי כמות הפיגועים בסדר עולה

- **אקספרסיביות:**

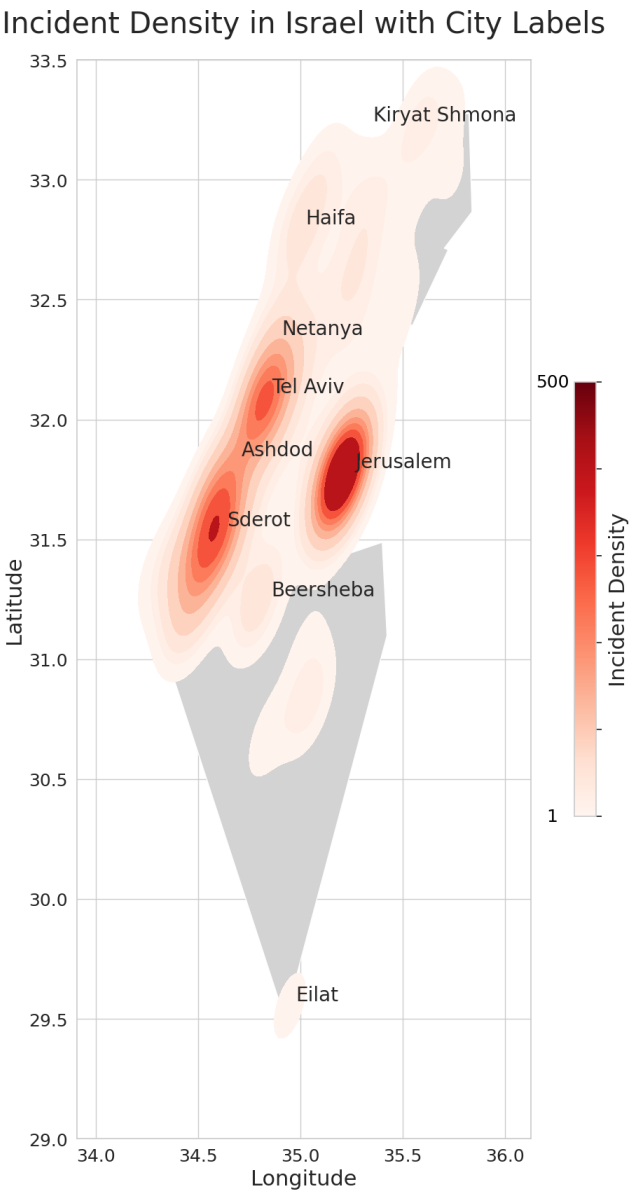
השימוש בגרף עמודות מתאים לסוג הנתונים, תכונה קטגוריאלית(עיר) ותכונה כמותית(תדירות התקיפות).

הערוץ הוויזואלי של אורך העמודות מייצג את התדירות של האירועים, נתון כמותי, וממוין לפי הכמות.

- **אפקטיביות:**

קל להשוות בין התדירות של הפיגועים בין הערים השונות באמצעות שימוש בערוצים יותר אפקטיביים. 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. בנוסף, סידור הערים לפי תדירות האירועים בסדר יורד מאפשר להבליט את הערים עם המספר הגבוה ביותר של אירועים, כמו ירושלים הממוקמת בראש הרשימה, ומשפר את יכולת ההבנה וההשוואה של הנתונים.

חלופה 2 למטלה 1:



- **Marks:**

  נקודות – כל נקודה מייצגת פיגוע

- **Channels:**

  Spatial Region – מיקום בציר האופקי והאנכי לפי נ"צ

  Color Saturation – הרוויה מייצג את כמות הפיגועים שהיו באותו אזור, ככל שכהה יותר ככה יש יותר פיגועים

  Area – משמש לייצוג רוחב אזורי החום, כאשר אזורים גדולים יותר מצביעים על אזורי התפשטות רחבים של אירועים.

- **אקספרסיביות:**

  הגרף מתאים לנתונים, שני תכונות כמותיות(קו אורך ורוחב) שמייצגות את הצירים ותכונה כמותית(תדירות התקיפות) שמייצגת את הרוויה

  הערוץ הוויזואלי של הרוויה מתאים והוא מייצג את תדירות התקיפות לפי הרוויה

- **אפקטיביות:**

  מפת החום עוזרת להבחין במקומות בהם יש ריכוז גבוה יותר של מקומות

  בחירת הרוויה בתור הרוויה הכהה לאזור שבו יש הרבה פיגועים מושכת את העין לעומת האזורים עם הרוויה הבהירה שבהם אין הרבה פיגועים

  תוויות הערים עוזר לזהות בערך את המיקום של הערים

- **<u>הצעה לשיפור:</u>** המיקום המרחבי הוא הערוץ היעיל ביותר עבור נתונים גיאוגרפיים, ומאפשר זיהוי מיידי של מיקום האירועים. בנוסף, הצבעים מסודרים באופן טבעי (גוונים בהירים מייצגים צפיפות נמוכה וכהים צפיפות גבוהה), מה שמקל על הבנת רמות הצפיפות. השילוב בין גווני הצבעים לתוויות הערים משפר את הנראות ומאפשר הבחנה קלה בין רמות צפיפות שונות, תוך זיהוי מוקדים מרכזיים כמו ירושלים ושדרות. השוני הברור בין גווני הצבעים תורם ליכולת הבחנה חדה בין אזורים צפופים יותר ופחות, ותוויות הערים מספקות הקשר ברור, המקל על הבנת מיקומי המוקדים ומידע הקשור בהם.

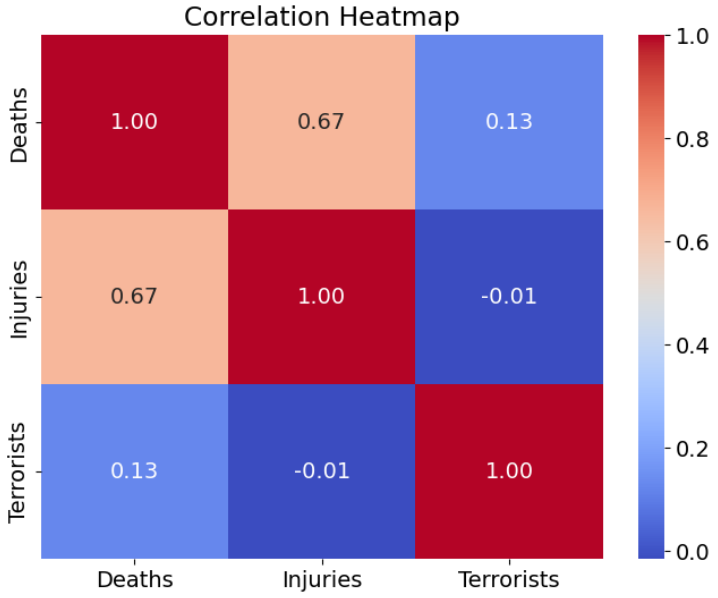
	חלופה 1 – גרף עמודות היסטוגרמה	חלופה 2 – גרף מפה של צפיפות
יתרונות	מהיר לקריאה, אפשר ישירות לדעת באיזו עיר יש הכי הרבה פיגועים. אפשר להעריך את הכמות של הפיגועים בכל עיר.	אפשר לזהות אזורים כללים בארץ שבהם יש הרבה פיגועים. זיהוי מהיר בעזרת הרוויה של מקומות שבהם יש הרבה פיגועים.
חסרונות	מוגבל רק לערים שמופיעות 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

2. <u>קורלציה בין מספר המחבלים למספר הנפגעים: האם קיים קשר בין כמות המחבלים באירוע לכמות הנפגעים?</u>

המטרה המרכזית של ניתוח הקורלציה היא לזהות קשרים בין המשתנים: מספר המחבלים (Terrorists), מספר ההרוגים (Deaths), ומספר הפצועים (Injuries). ניתוח זה מאפשר להבין האם קיים קשר ליניארי בין היקף האירוע (מספר המחבלים) לבין התוצאה האנושית (הרוגים ופצועים).

<u>חלופה 1 למטלה 2:</u>



- **Marks:**

כל תא במפת החום הוא סימון (Mark) שמציג את הקורלציה בין שני משתנים: מספר מחבלים, מספר הרוגים, ומספר פצועים.

- **<u>הצעה לשיפור:</u>**
- ריבועים - מייצגים את הקשרים בין משתנים.
- טקסט בתוך הריבוע – מייצג את ערכי הקורלציה בין 2 משתנים
- קווי רשת – תוחמים את הקורלציות השונות ומבדילים בין הקטגוריות.
- **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) /*המיקום של כל ריבוע במטריצה תואם את המשתנים שאותם הוא משווה) לדוגמה "Deaths", מול."Injuries" */

מספרים על התאים - ערך הקורלציה המספרי המדויק בין המשתנים /* המספרים בריבועים מספקים ערך מספרי מדויק של הקורלציה, משלים את ערוץ הצבע/*.

- **אקספרסיביות:**

מציגה עוצמת קשר ליניארי בלבד: הערכים והצבעים מתארים את עוצמת הקורלציה בלבד, מבלי

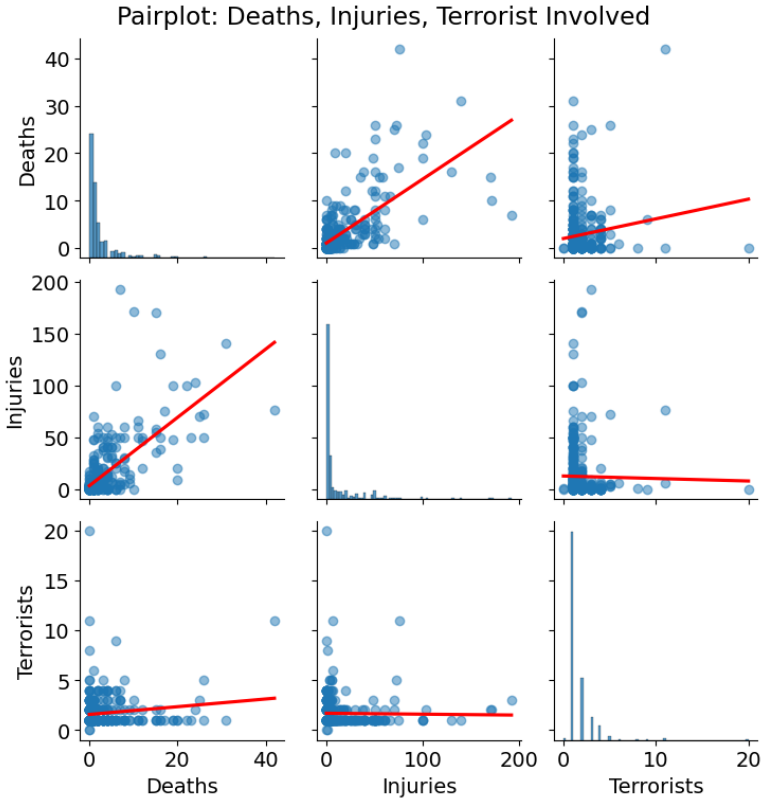
לפרט על צורת הקשר, על חריגות או על פיזור התצפיות. זה מגביל את המידע שאפשר להפיק מהמפה.

- **<u>הצעה לשיפור:</u>** הנתונים הרציפים (Ordered Data) מוצגים בסולם צבע רציף (Magnitude Channel), שמתאים להבעת ערכים כמותיים בצורה ברורה. המיקום המרחבי של המשתנים בצירי המטריצה מבטיח ייצוג עקבי וברור, והמספרים המוצגים בתוך הריבועים מוסיפים רמת דיוק המאפשרת קריאה ישירה של ערכי הקורלציה. בנוסף, השימוש בצבעים ובסולם מונע פרשנויות שגויות, מדגיש רק את המידע הרלוונטי, ומספק ייצוג חזותי תואם ואמין של הקשרים בין המשתנים.
- **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.

- **<u>הצעה לשיפור:</u>** השימוש בטקסט בתוך הריבועים מאפשר קריאה מדויקת של ערכי הקורלציה, בעוד הצבעים עוזרים לתפיסה מהירה של עוצמת הקשרים. הסולם הצבעוני יוצר הבחנה ברורה בין ערכים חיוביים, שליליים ואפסיים, מה שמקל על זיהוי קשרים חזקים או חלשים. הצבעים הבולטים, כמו אדום וכחול כהה, מושכים את העין למקרים של קשרים חזקים. בנוסף, הטקסט והמבנה המסודר של המטריצה תורמים לקריאות ומאפשרים הבנה קלה של הקשרים בין המשתנים.

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.

- **לשקול להוסיף:** הנקודות כולן בצבע אחיד (כחול), בעוד קו הרגרסיה (אדום) מדגיש את מגמת הקשר.

- **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.

לשקול להוסיף: קווי הרגרסיה האדומים מוסיפים רובד חזותי ברור המדגיש את כיוון ועוצמת הקשר בין זוגות המשתנים, ומסייעים בזיהוי מגמות מרכזיות. בנוסף, הוויזואליזציה מתמקדת במידע הרלוונטי בלבד ומציגה את כל הקשרים האפשריים בין המשתנים, מבלי לכלול מידע שאינו נתמך בנתונים, מה שמבטיח דיוק ואמינות.

- **אפקטיביות:**

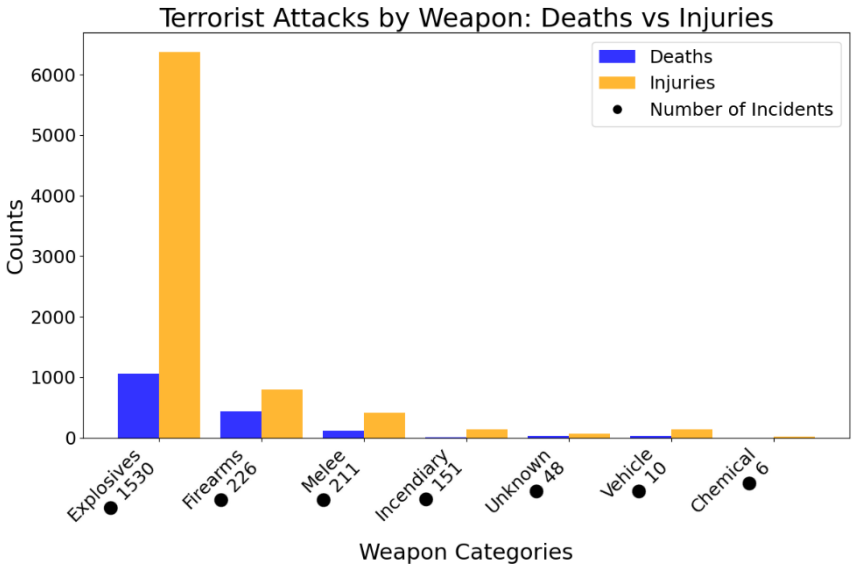
מתאימה לניתוח מעמיק של קשרים ומגמות: מטריצת גרף פיזור מאפשרת להבין לא רק האם קיים קשר בין שני משתנים, אלא גם את אופי הקשר (ליניארי, לא ליניארי) ואת פיזור הנתונים. בעזרת קווי המגמה, ניתן לזהות את כיוון הקשר ולזהות אם קיימות קבוצות חריגות שדורשות בחינה נוספת.

	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

3. <u>מהם ההבדלים בין סוגי הנשק בתרומתם לחומרת האירועים(מספר נפגעים והרוגים) ותדירות האירועים מכל סוג נשק?</u>
נרצה להשוות בין סוגי נשק שונים שבהם המחבלים השתמשו בכל פיגוע, בהתחשבות במספר הנפגעים וההרוגים וכמות הפיגועים שבוצעו בכל סוג נשק. נרצה גם להבחין בסוגי נשק קטלניים יותר(גורמי נזק).

<u>חלופה 1 למטלה 3:</u>



- **Marks:**

עמודות: מייצגות את מספר הנפגעים (Injuries) וההרוגים (Fatalities)לכל סוג נשק.

נקודות שחורות: מתחת לכל קטגוריה מצוינת כמות הפיגועים (מספר האירועים) באמצעות נקודה ומספר.

- **Channels:**

מבנה הגרף מבוסס על מספר מרכיבים עיקריים: במישור האופקי (ציר ה-X), המיקום משקף את סוג הנשק, בעוד שבמישור האנכי (ציר ה-Y) מוצגת כמות ההרוגים והפצועים עבור כל קטגוריה. העמודות הכחולות מייצגות את מספר ההרוגים (fatalities), ואילו העמודות הכתומות מציגות את מספר הפצועים (injuries). מתחת לציר X הוספו נקודות שחורות עם מספרים המציינות את מספר האירועים הכולל בכל קטגוריה, ובכך מספקות מידע משלים חשוב להבנת היקף התופעה.

- **<u>הצעה לשיפור:</u>** המיון מדגיש את סוגי הנשק שנמצאים בשימוש הנפוץ ביותר, ומאפשר השוואה מהירה וברורה של סוגי הנשק המשפיעים ביותר מבחינת תדירות השימוש.
- המיון לפי מספר האירועים מסייע למשתמש לזהות את הקשר בין תדירות השימוש לבין מספר ההרוגים והפצועים.
- **אקספרסיביות:**

הגרף מציג את הנתונים על פיגועי טרור לפי סוג הנשק שבו נעשה שימוש. כל עמודה מייצגת קטגוריה של סוג נשק, כאשר העמודות הכחולות מסמלות את כמות ההרוגים והעמודות הכתומות את כמות הפצועים. בנוסף, מתחת לכל קטגוריה מופיעה נקודה שחורה עם מספר האירועים עבור אותה קטגוריה, מה שמספק מידע נוסף על כמות הפיגועים. המיקום בציר X מייצג את סוג הנשק, בעוד המיקום בציר Y מייצג את כמות ההרוגים והפצועים. השימוש בצבעים שונים לעמודות מאפשר הבחנה ברורה בין נתוני ההרוגים והפצועים, והנקודות השחורות מתחת לציר X מוסיפות רובד נוסף של מידע בצורה קריאה וברורה. הגרף מאפשר לזהות בקלות את קטגוריית הנשק בעלת ההשפעה הגדולה ביותר, כמו Explosives, ואת תרומתה מבחינת כמות ההרוגים והפצועים, תוך שמירה על נראות אסתטית ונוחות קריאה.

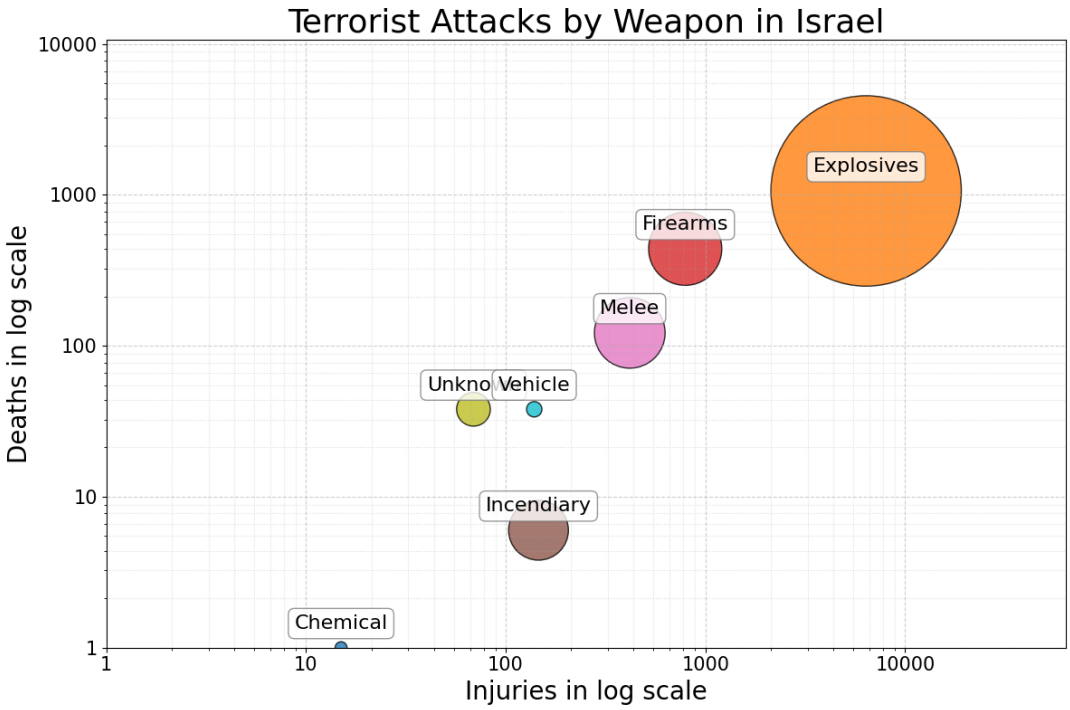
- **אפקטיביות:**

הגרף מאפשר הסקת מספר תובנות והשוואות משמעותיות: ניתן לזהות בקלות את הקטגוריה

המובילה במספר הפגיעות (Explosives) ולהבין את מידת השפעתה מבחינת הרוגים ופצועים. הקטגוריות בעלות מספר אירועים נמוך יותר מודגשות באופן ייחודי באמצעות נקודות קטנות מתחת לעמודות. מבחינת בחירת הצבעים, השימוש בכחול וכתום עבור קטגוריות ההרוגים והפצועים יוצר ניגודיות ברורה ומאפשר קריאות נוחה של הנתונים. תוספת הנקודות השחורות מתחת לציר X מהווה פתרון לייצוג מידע נוסף על מספר האירועים בצורה חזותית וקריאה בבירור.

- **<u>לשקול להוסיף:</u>** סידור העמודות לפי תדירות האירועים מדגיש את סוגי הנשק הנפוצים ביותר, כמו Explosives, שגורמים למספר גבוה במיוחד של נפגעים ותדירות גבוהה של שימוש.

<u>חלופה 2 למטלה 3:</u>



- **Marks:**

עיגולים :כל בועה מייצגת קטגוריה של סוג הנשק שבו נעשה שימוש בפיגועי טרור.

- **<u>הצעה לשיפור</u>**:

גודל העיגול: מספר האיורעים בהם נעשה שימוש בנשק.

מיקום: מייצג את כמו תההורגים והפצועים לכל סוג נשק.

טקטס: תוויות המציינות את סוגי הנשק עבור כל עיגול.

- **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. גודל 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.

- **<u>הצעה לשיפור:</u>** הוויזואליזציה מדגישה בצורה ברורה את המגמות בשימוש בסוגי נשק שונים בפיגועי טרור בישראל באמצעות מיקום עיגולים בסקאלה לוגריתמית, המשקפת את כמות ההרוגים והפצועים. השימוש בסקאלה לוגריתמית מאפשר דיוק גבוה להצגת ערכים בטווחים רחבים, וגודל העיגולים מייצג את תדירות האירועים עבור כל סוג נשק, מה שמוסיף ממד נוסף להבנת הנתונים. צבעי העיגולים ותוויות השמות מאפשרים הבחנה ברורה בין הקטגוריות ומספקים זיהוי מיידי של סוגי הנשק. הוויזואליזציה מבליטה מגמות מרכזיות, כמו ההשפעה הגבוהה של Explosives מבחינת מספר הנפגעים ותדירות השימוש, לעומת נשקים כמו Chemical שלהם השפעה ותדירות נמוכה. עיצוב זה מסייע לזהות בקלות את הדפוסים המרכזיים ולהבין את השפעת כל סוג נשק על היקף ותוצאות הפיגועים.

	חלופה 1 – **גרף בועות**	**חלופה 2 –** **גרף עמודות**
יתרונות	גרף הבועות מאפשר הצגת קשר בין שני משתנים כמותיים, כמו פצועים והרוגים, באמצעות מיקום בצירים, ובכך מספק תמונה ברורה של היחסים בין הקטגוריות. גודל הבועה מעניק דרך אינטואיטיבית להראות את מספר הפיגועים לכל קטגוריה, מה שמקל על זיהוי קטגוריות עם השפעה רבה יותר. תוויות הבועות מספקות זיהוי ברור של הקטגוריות באופן ישיר, ללא צורך באגדה נפרדת, והגרף מתאים במיוחד להצגת נתונים יחסיים ולבחינת קשרים בין משתנים.	גרף העמודות מאפשר קריאה והשוואה פשוטה של כמות ההרוגים והפצועים בין הקטגוריות בצורה מובנית וברורה. השימוש בצבעים, כמו כחול להרוגים וכתום לפצועים, מבליט את ההבדלים בצורה נוחה להבנה. בנוסף, הנקודות השחורות מתחת לכל עמודה עם מספר הפיגועים מספקות מידע נוסף בצורה ישירה ואסתטית. גרף זה מתאים במיוחד למשתמשים שמעדיפים השוואה פשוטה וברורה בין קטגוריות שונות, כמו גם לקריאה מהירה של הנתונים.

חסרונות	גרף הבועות עלול להיות פחות קריא כאשר קיימות קטגוריות רבות, מה שעלול לגרום לצפיפות יתר בגרף. קטגוריות קטנות עם מספר אירועים קטן עלולות להיראות זניחות בשל גודל הבועה הקטן שלהן. בנוסף, התלות בציר לוגריתמי עשויה להקשות על הבנת הנתונים למי שאינו רגיל לעבוד עם סולמות לוגריתמיים. הגרף פחות מתאים להצגת נתונים אבסולוטיים כאשר יש צורך בהשוואה ישירה בין קטגוריות.	גרף העמודות אינו מציג קשרים בין משתנים (לדוגמה, בין פצועים להרוגים), אלא מתמקד בהשוואה קטגורית בלבד. עומס המידע, הנובע משילוב עמודות, נקודות שחורות ומספרים, עלול להכביד על המשתמש בהשוואה פשוטה ומהירה. קטגוריות עם מספר אירועים נמוך עלולות להיראות זניחות, במיוחד אם העמודות שלהן קטנות מאוד ביחס לקטגוריות בולטות יותר, מה שעלול לטשטש את המשמעות שלהן בתוך הגרף.

הוויזואליזציה שנבחרה היא: 2

הסבר על העיצובים שנבחרו ועל יישומם

עיבוד מוקדם של הנתונים:

0. בפרויקט הורדנו את הנתונים שלנו מאתר קאגל
1. Data Filter: סיננו את אירועי הטרור מישראל (קוד ב-R ב-git)
2. Data Cleaning: לאחר מכן ביצענו ניקוי של הערכים השליליים בכך שהפכנו אותם ל-Null (קובץ ipynb ב-git)

האם נאלצנו לשנות את מימוש הויזואליזציה:

במהלך הפרויקט שינינו את המימוש של העיצוב של מטלה 1, בחרנו לעשות מפת חום ובהתחלה המפה לא הייתה אינטראקטיבית(כמו שתכננו בחלופה), לאחר מכן הכרנו את ספריית Folium שעזרה לנו לייצר מפה אינטראקטיבית ובחרנו לעצב ככה את המפה(היא עדיין מראה אותו הדבר).
ההבדל בין העיצובים הוא האינטראקטיביות של הויזואליזציה.

צילומי מסך של הדאש בוארד:

ב-Dashboard אנו מציגים את השאלות ואת הגרפים כאשר בדף הראשון יש Data Overview ובכל דף אחר יש מטלה, השתדלנו לשמור עליו כמה שיותר נקי שיהיה מינימליסטי אבל שיכיל את כל הכלים והמלל שצריך. לצד כל ויזואליזציה יש הוראות שימוש/קריאה שלה ובנוסף יש תיאור המטלה שעליה אנו עונים באמצעות הויזואליזציה.

מבט כללי על הנתונים:

Total Rows
2,182

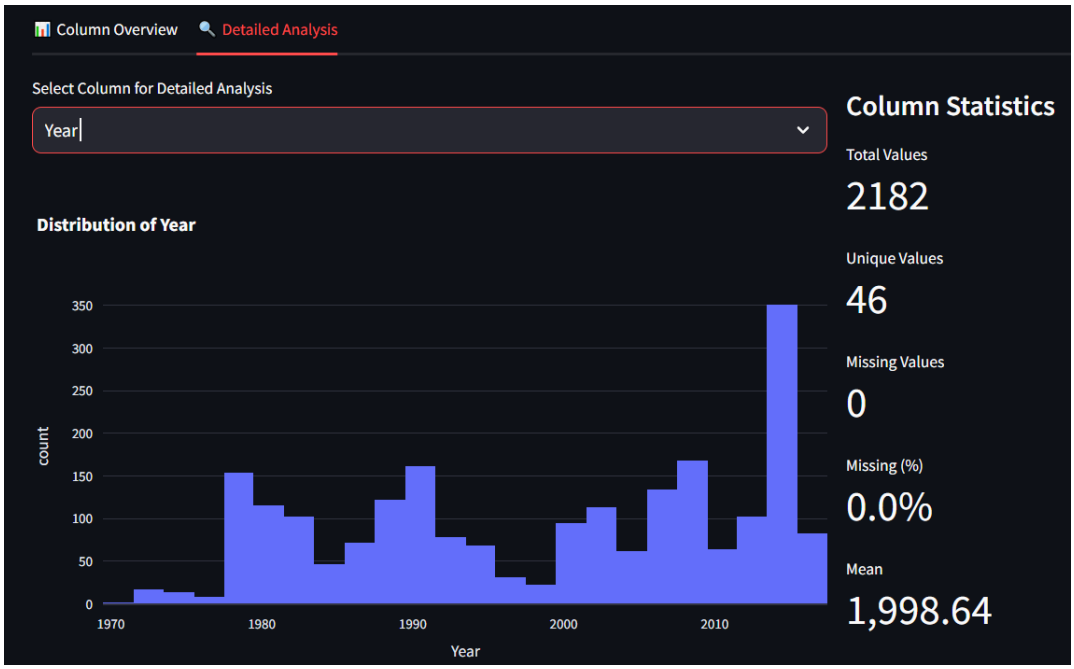
Total Columns
10

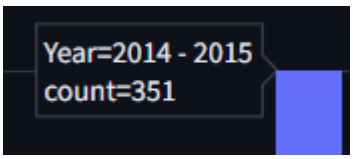
Missing Values
108

📊 Column Overview 🔍 Detailed Analysis

	Column	Type	Unique Values	Missing Values (%)	Min	Max	Mean	Description
0	Year	int64	46	0.0%	1,971.00	2,017.00	1,998.64	The year the event occurred
1	Month	int64	12	0.0%	1.00	12.00	6.40	The month the event occurred
2	Day	int64	32	0.0%	0.00	31.00	15.77	The day the event occurred
3	City	category	367	0.0%	N/A	N/A	N/A	The name of the city village or town where the event occurred
4	Latitude	float64	416	2.5%	29.55	33.42	31.84	Latitude
5	Longitude	float64	420	2.5%	34.17	35.86	34.93	Longitude
6	Terrorists	int64	12	0.0%	0.00	30.00	0.33	The total number of terrorists participating in the event
7	Deaths	int64	29	0.0%	0.00	42.00	0.78	The total number of deaths in the event
8	Injuries	int64	69	0.0%	0.00	192.00	3.64	The number of wounded
9	Weapon Type	category	8	0.0%	N/A	N/A	N/A	The type of weapon used in the event

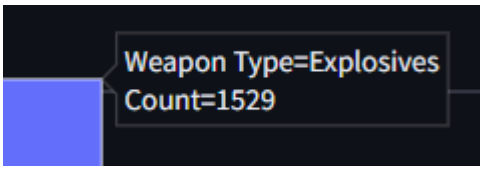
אפשר לבחור פיצ'ר מסוים
ולראות נתונים סטטיסטים
על אותו הפיצ'ר:





אם שמים את העכבר מעל אפשר
לראות את הכמות המדויקת:

עבור עמודות קטגוריאליות זה נראה כך (אם יש יותר מ-10
ערכים אפשריים, נראה רק את ה-10 הנפוצים ביותר):



אם שמים את העכבר מעל אפשר
לראות את הכמות המדויקת:

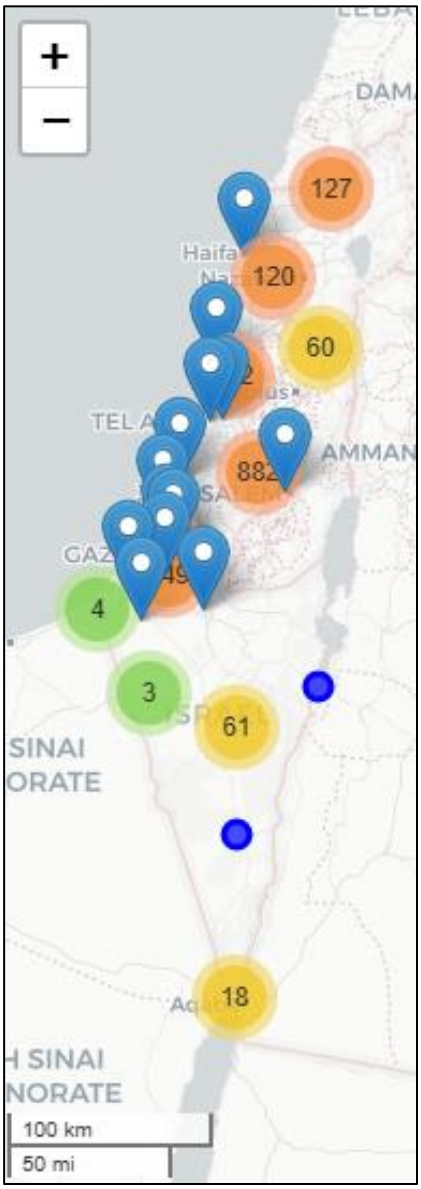
מטלה 1

הסבר על המפה:

- הנקודות הכחולות מייצגות את מיקומי תקיפות הטרור
- צבע הנקודה והמספר שמופיע בנקודות מייצג את מספר אירועי הטרור באותו מיקום
- לחיצה על נקודה תציג פרטים נוספים על האירוע, כולל שם העיר, מספר הנפגעים והמיקום המדויק

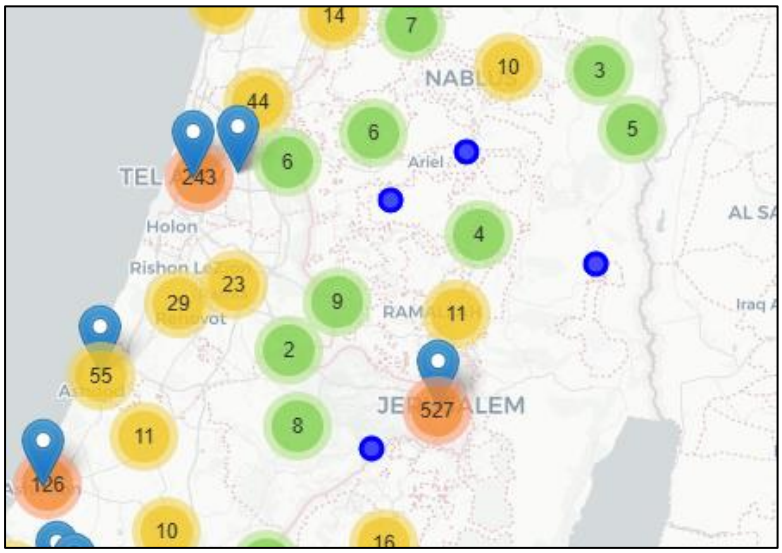
מקרא:

- כחול - אירוע 1
- ירוק - 2 עד 10 אירועים
- צהוב - 11 עד 100 אירועים
- כתום - יותר מ-100 אירועים
- - ערים מרכזיות בישראל שבהן התרחשו הכי הרבה אירועי טרור



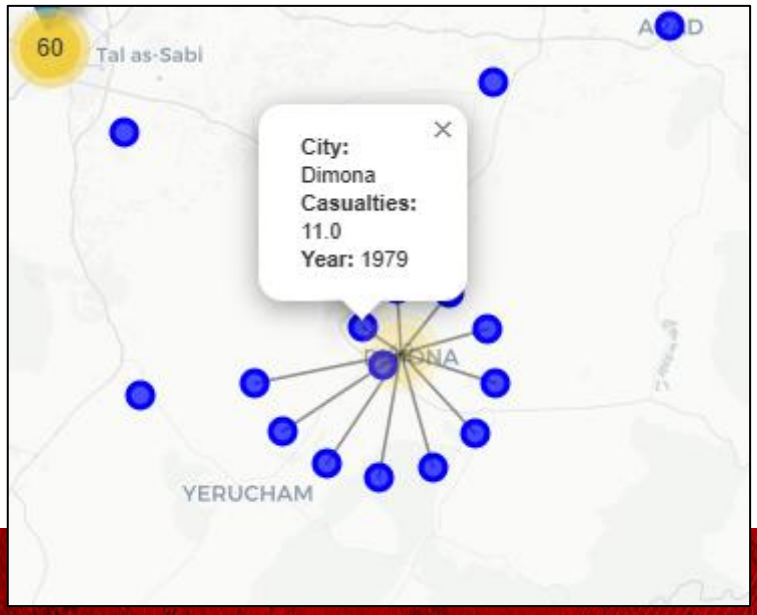
לאחר לחיצה על הפלוס נוכל להתקרב:

דוגמה ל-Navigate: Zoom



כשמתקרבים יותר לנקודה ולוחצים עליה:

דוגמה ל-ToolTip



מטלה 2

בהתחלה נבחר את כל ציר הזמן:

מטריצת הגרפים מציגה את כל הקשרים האפשריים בין המשתנים הבאים:

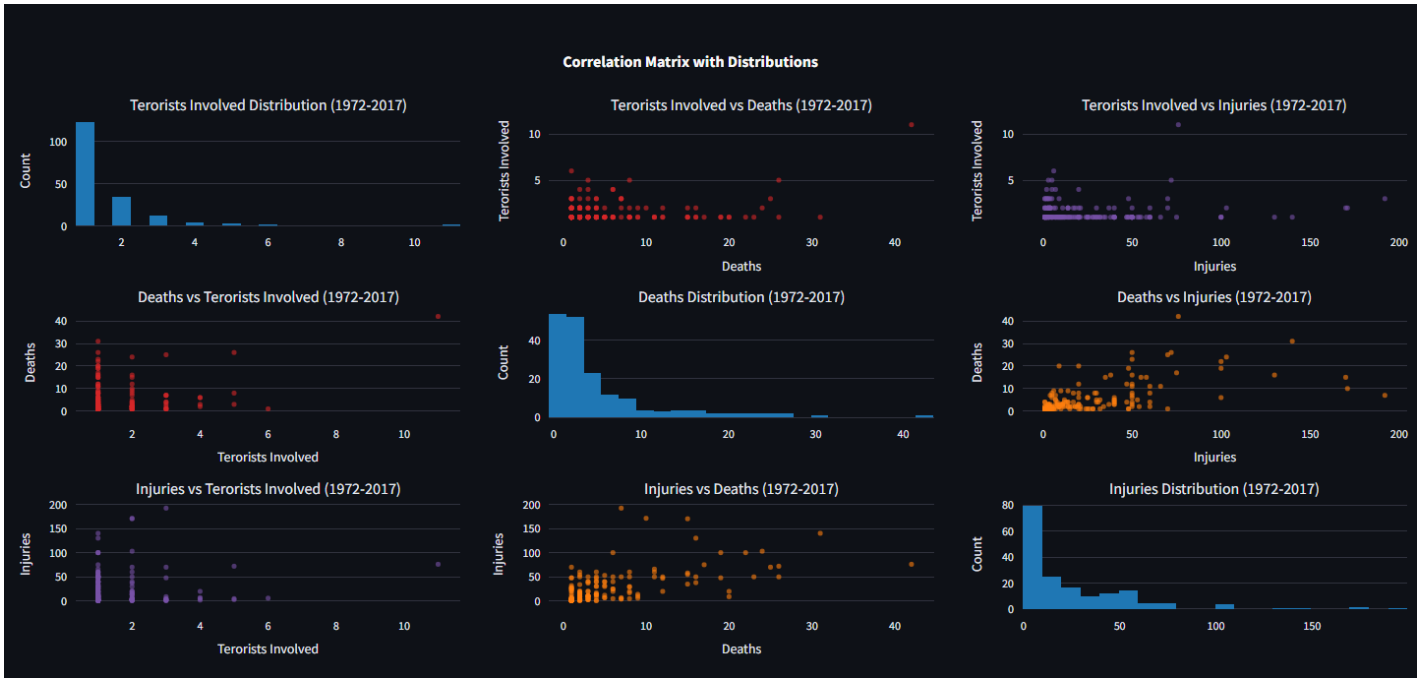
- מספר המחבלים (Terorists Involved)
- מספר ההרוגים (Deaths)
- מספר הפצועים (Injuries)

כל משבצת במטריצה מציגה את הקשר בין שני משתנים, כאשר:

- באלכסון - התפלגות המשתנה
- מעל/מתחת לאלכסון - פיזור הנקודות בין שני המשתנים

Select Year Range

1972 — 2017



כששמים את העכבר מעל נקודה מסוימת:

דוגמה ל-ToolTip

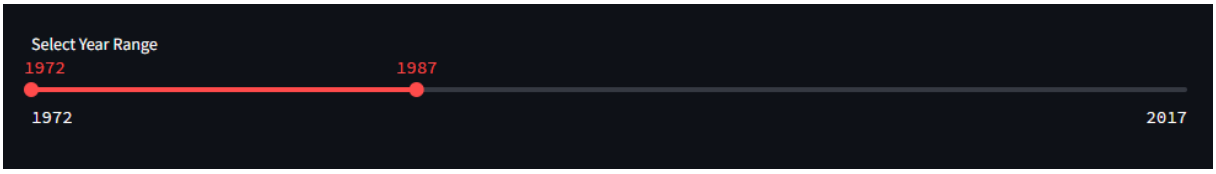
מטריצת קורלציות:

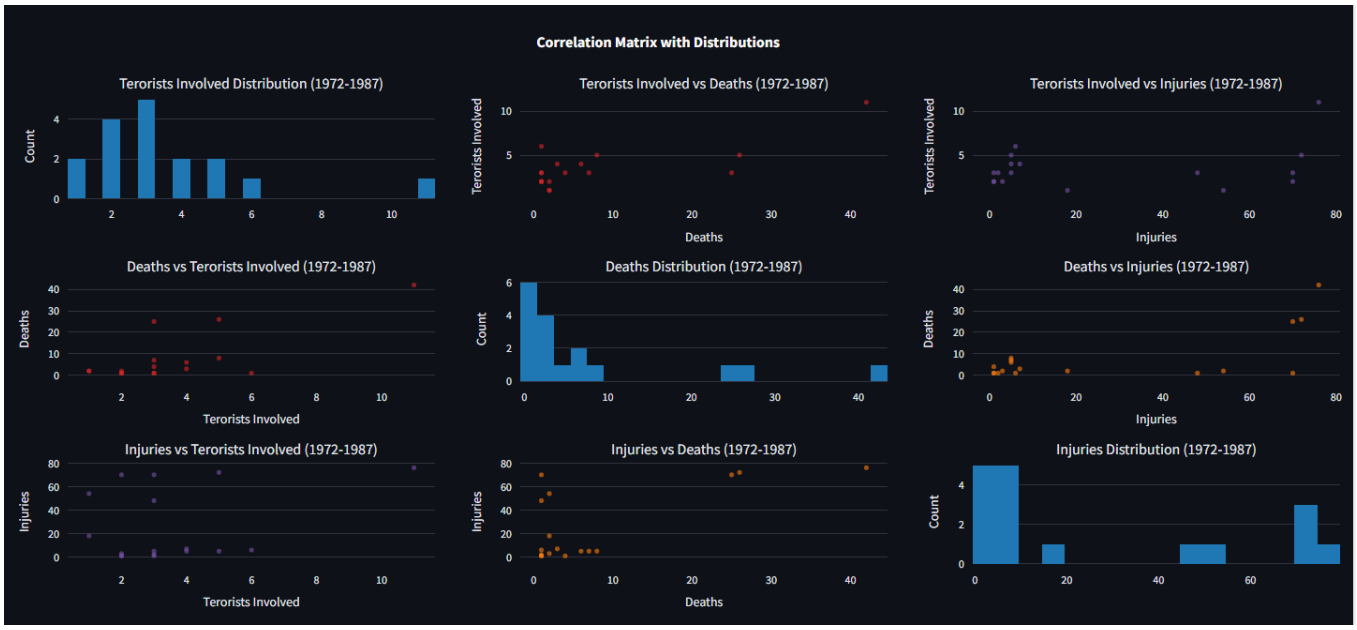
Correlation Matrix

	Terrorists	Deaths	Injuries
Terrorists	1.000	0.292	0.051
Deaths	0.292	1.000	0.592
Injuries	0.051	0.592	1.000

- 1: Perfect positive correlation
- 0: No correlation
- -1: Perfect negative correlation

נבחר ציר זמן קצר





המטריצה התעדכנה בהתאם

Correlation Matrix

	Terrorists	Deaths	Injuries
Terrorists	1.000	0.744	0.292
Deaths	0.744	1.000	0.653
Injuries	0.292	0.653	1.000

מטלה 3

ניתוח התפתחות הטרור לאורך זמן:

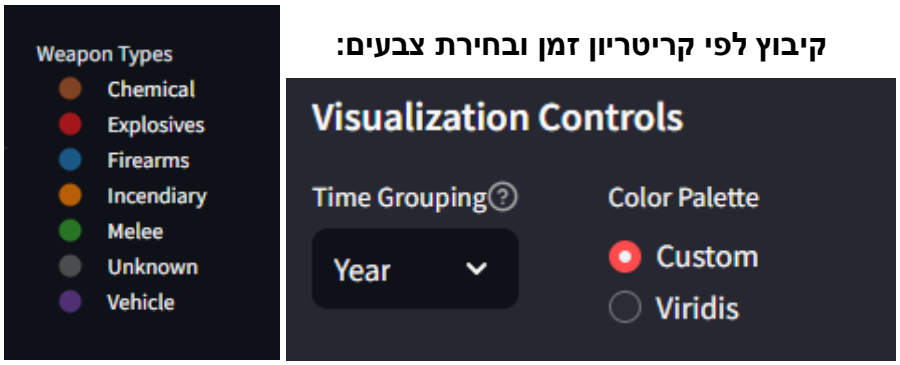
הויזואליזציה מציגה את התפתחות אירועי הטרור לאורך זמן, כאשר:

- כל בועה מייצגת סוג נשק
- גודל הבועה מייצג את מספר האירועים
- מיקום הבועה מראה את היחס בין מספר ההרוגים והפצועים
- הצבע מבדיל בין סוגי הנשק השונים
- האנימציה מראה את השינוי לאורך זמן

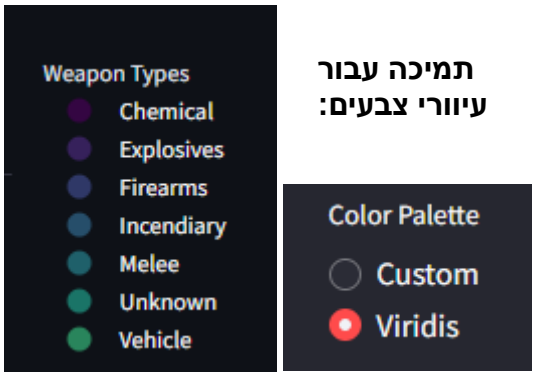
ניתן לשלוט בויזואליזציה באמצעות:

- כפתורי הפעלה/עצירה
- בחירת רמת הקיבוץ בזמן (חודשי/רבעוני/שנתי)
- סף מינימלי למספר האירועים

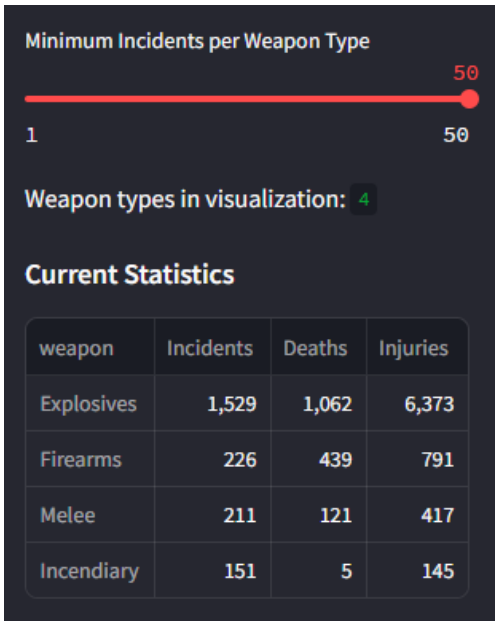
קיבוץ לפי קריטריון זמן ובחירת צבעים:



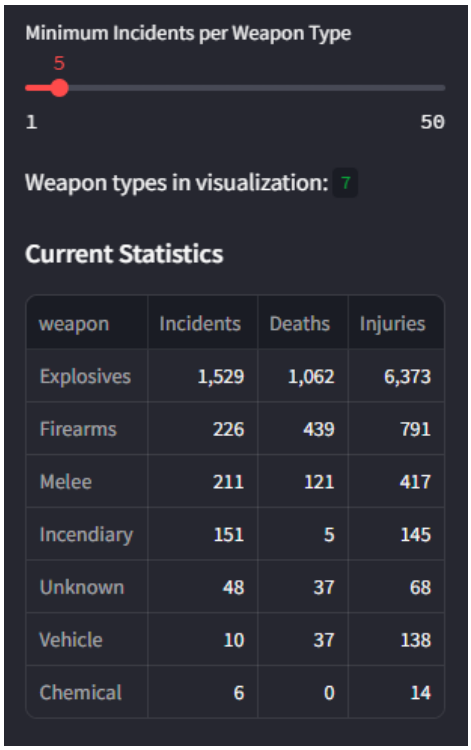
תמיכה עבור עיוורי צבעים:

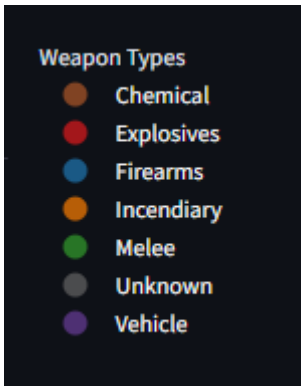


אפשר לסנן פריטים באמצעות מספר מינימלי של אירועים - מינימום 50:

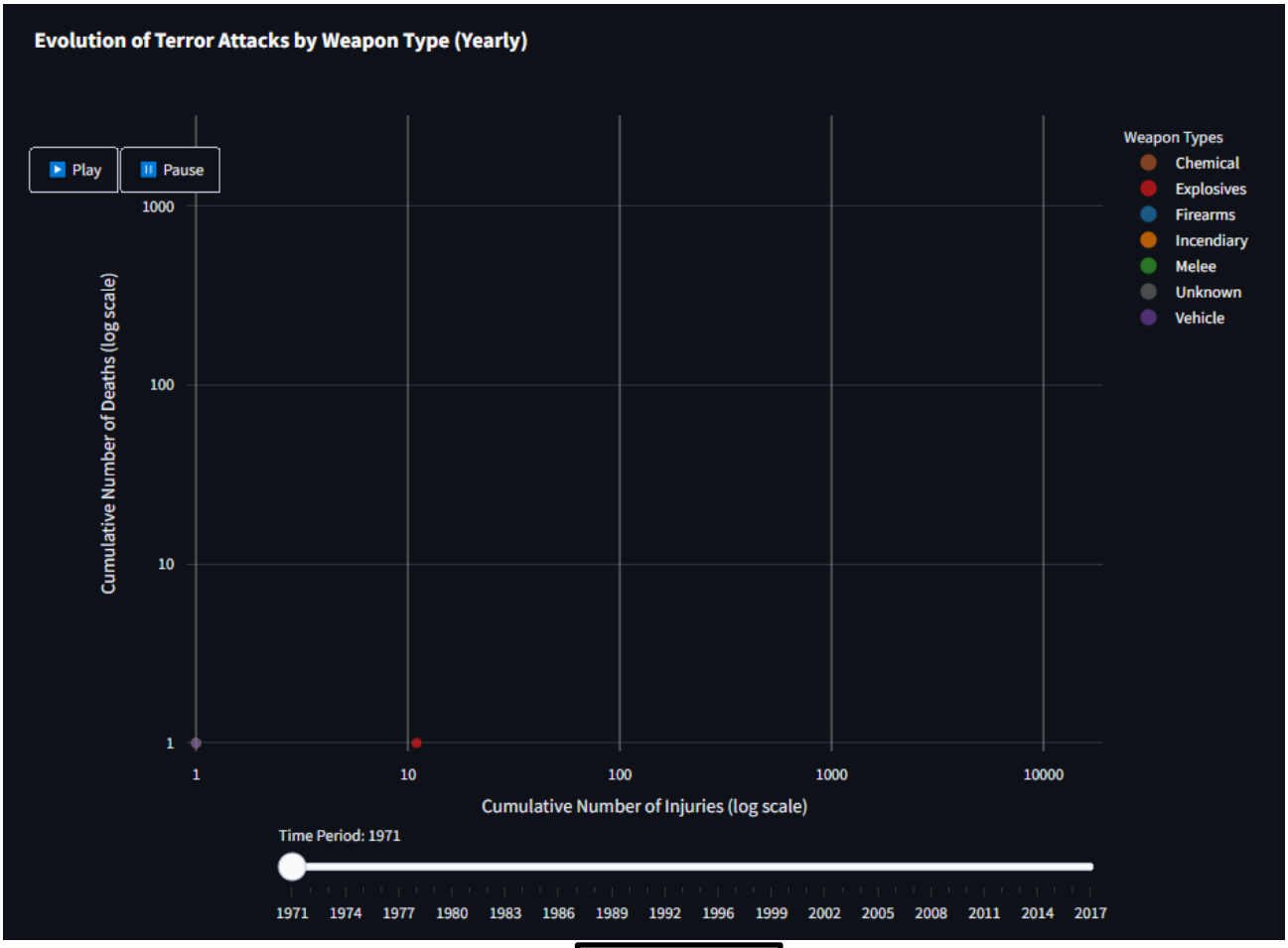


אפשר לסנן פריטים באמצעות מספר מינימלי של אירועים - האנימציה בהתחלה:

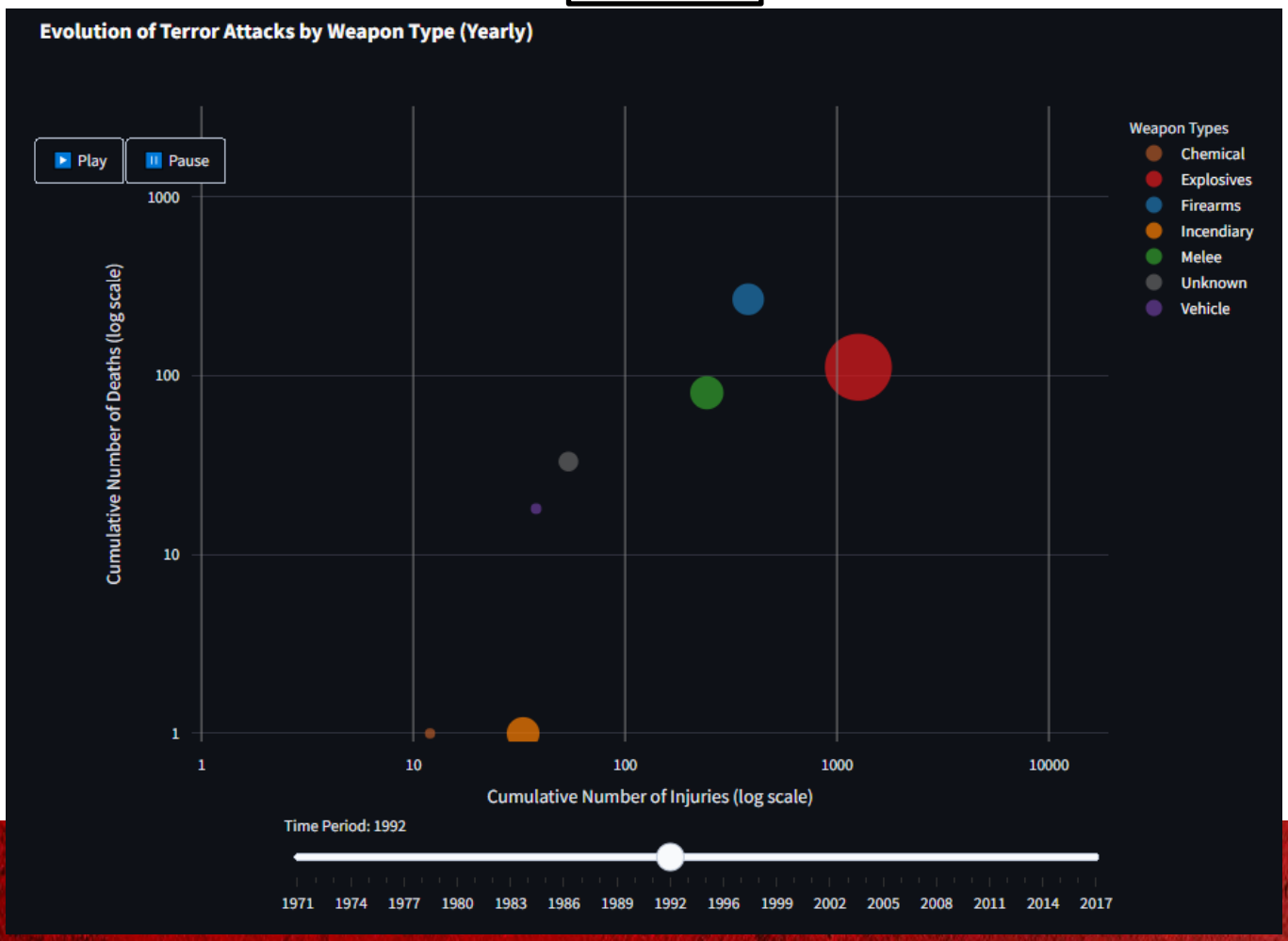




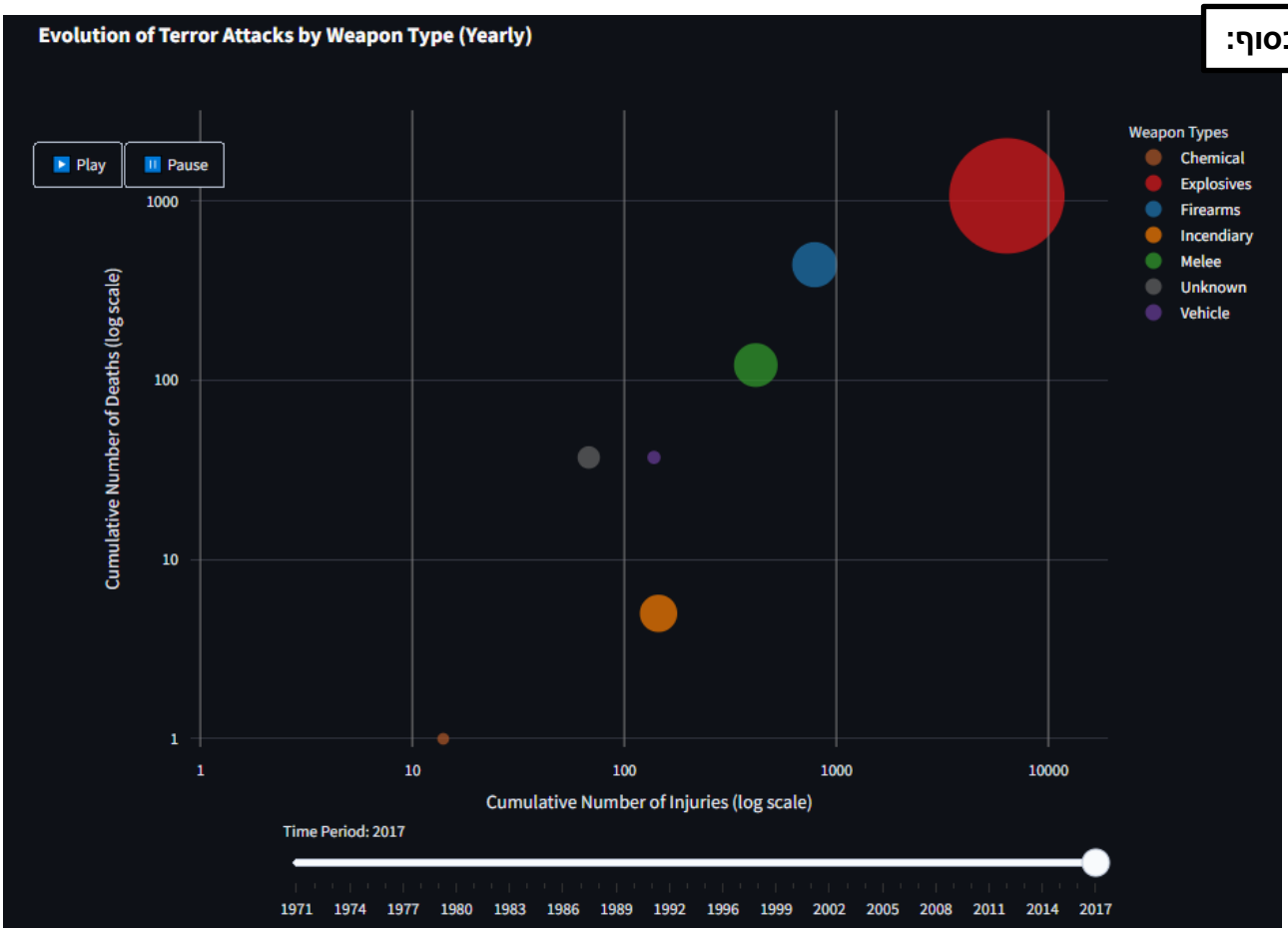
האנימציה בהתחלה:



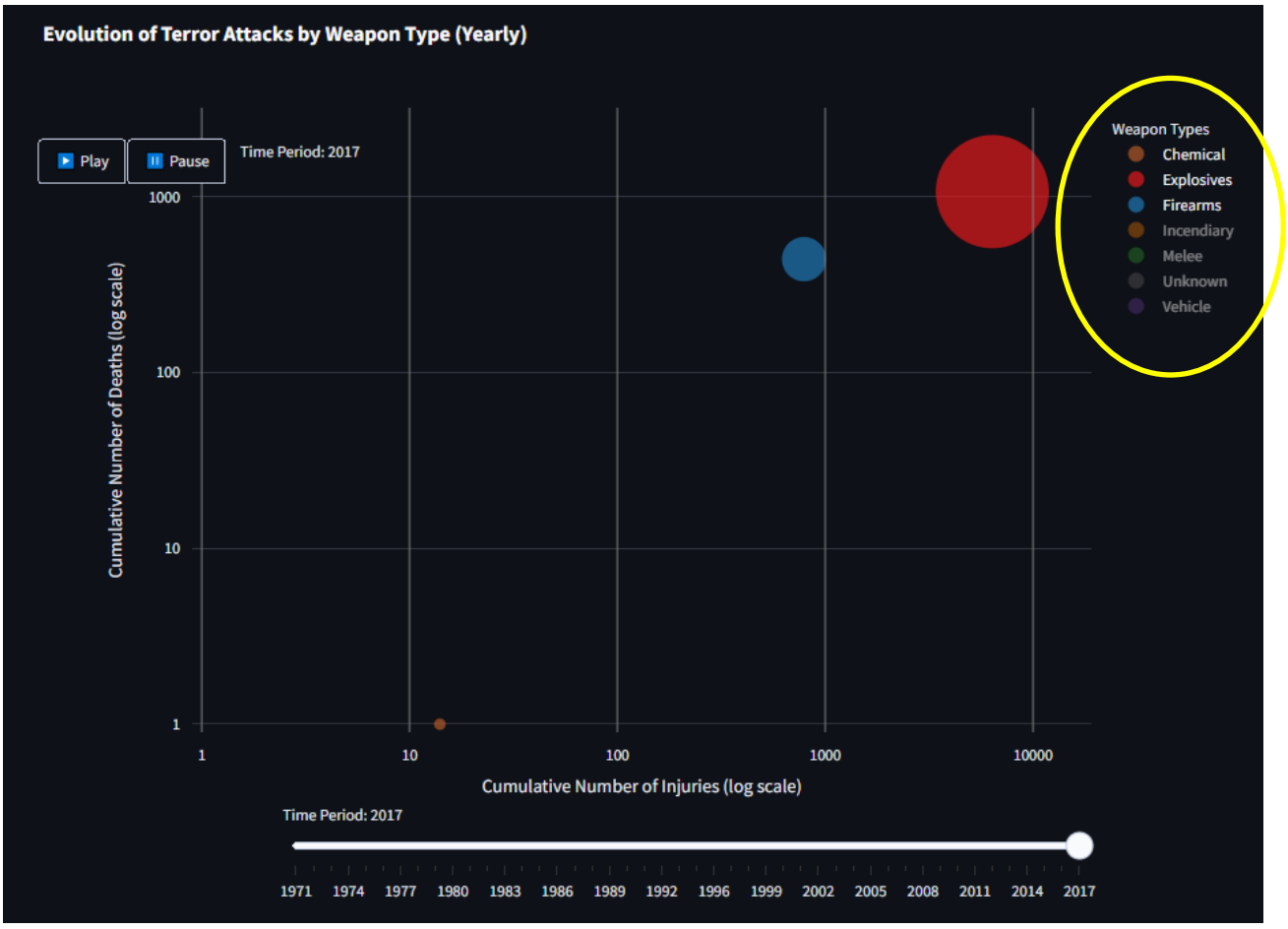
האנימציה באמצע לאחר שלחצנו Play: **Aggregate**



האנימציה בסוף:

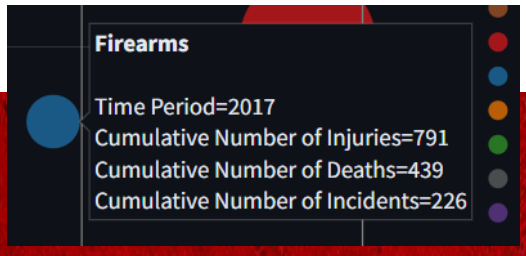


אפשר לסנן פריטים ידנית ולא לראות אותם:



כששמים את העכבר מעל נקודה מסוימת:

דוגמה ל-ToolTip



הסבר קצר מה עשינו בקוד:

השתמשנו בספריית Pandas לעיבוד הנתונים ולביצוע ויזואליזציות השתמשנו בספריות Plotly ו-Folium, בשביל לממש את ה-Dashboard השתמשנו בספריית Streamlit.

לקחנו מאתר Streamlit קוד של Template של DashBoard גנרי(על GDP), לאחר מכן קראנו את הקוד שהיה, הבנו את ה-Syntax ומחקנו את כל הקוד ממנו בשביל שיהיה לנו אפליקציה נקיה.

בדף הבית התבססנו על הקוד שלנו בחלק ניקוי הנתונים והצגנו אותו בטבלה עם HTML.

במטלה הראשונה נעזרנו בדוקומנטציה של ספריית Folium, לא היה הרבה מה לשנות, הספריה די פשוטה, שינינו לפי מה הוא יקבע את המיקום, הגדרנו מה יופיע ב-popup והגדרנו גם נ"צ של ערים עם מספר גבוה של פיגועים.

במטלה השנייה התבססנו על הקוד שלנו מהעיצובים החלופיים, אז השתמשנו בספריית matplotlib והמרנו את הקוד לספריית Plotly. הוספנו אליו גם קוד HTML כמו hover popup ובטבלת קורלציות.

במטלה השלישית התבססנו על קוד מתוך ה-Dash_Example.ipynb שהעלתם אלינו למודל, בגרף Total proportion infected by average salary בשילוב עם קוד מהדוקומנטציה של הספריה, בסוף יותר השתמשנו בקוד מהדוקומנטציה כי היה שם יותר אפשרויות לשנות דברים, בחרנו מי יהיה הערך שיופיע בתור כל בועה, הגדרנו את הצירים ובחרנו מיקומים לכל הכפתורים(הפעלה/עצירה) ובנוסף גם לטקסטים שקיימים ול-Slider של ציר הזמן.

בנוסף נעזרנו במהלך כתיבות הקוד עם Chat GPT בטיפול בשגיאות בקוד ב-Streamlit ושינוים באינטראקציות(כמו הגדרת גודל פונט ב-ToolTip)