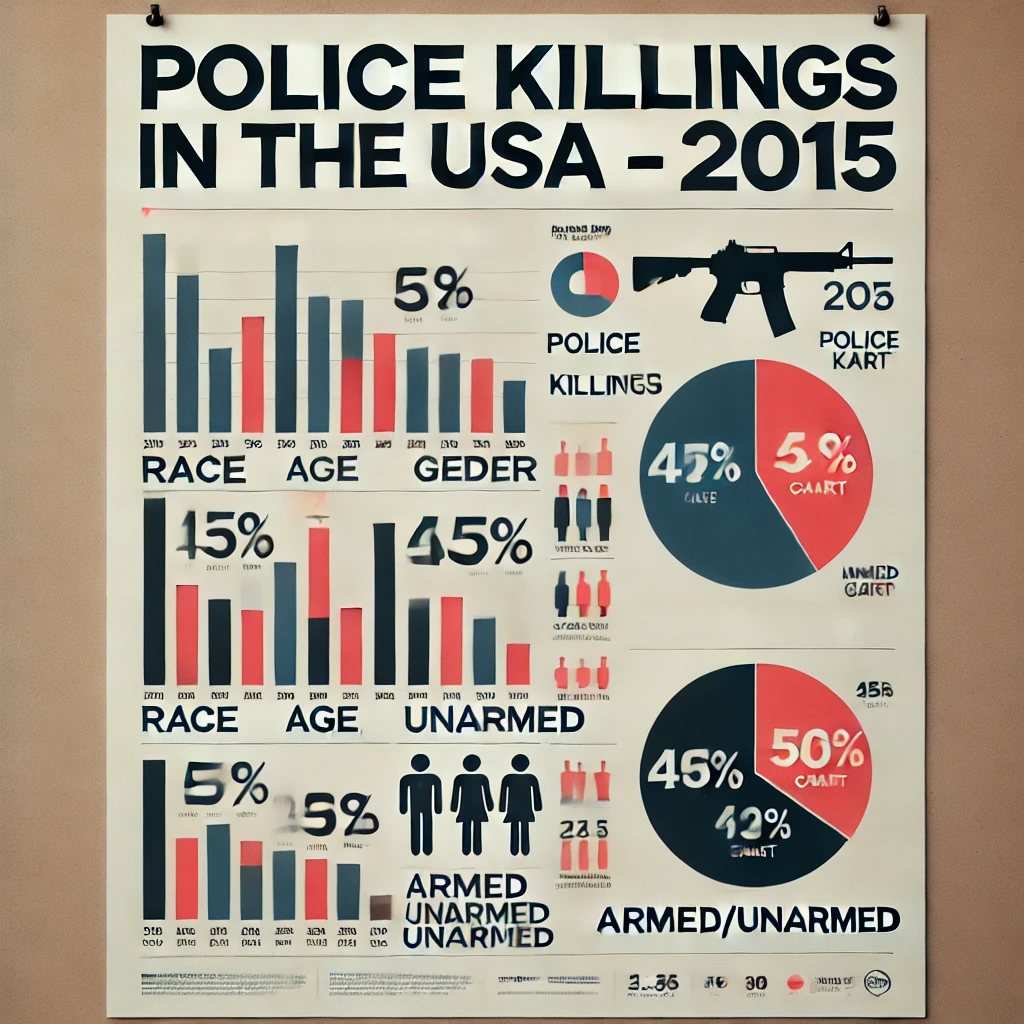
**מציאת הדפוסים והמגמות הכוללות של הרציחות המשטרתיות ברחבי ארה"ב ב-2015**

קורס - ויזואליזציה של מידע

תאריך הגשה – 20/07/2024

**מגישים - נופר סלוק 318502721 , רועי גרבר 209466614**



ניתוח נתונים ומטלות-

מטרת הפרויקט שלנו היא לנתח את הדפוסים והמגמות של רציחות משטרתיות ברחבי ארצות הברית בשנת 2015, תוך התמקדות בהבנת השפעת גורמים דמוגרפיים, גיאוגרפיים וסוציו אקונומיים. הבנת חשיבות נושאים אלו יכולה לתמוך בכמה תחומים ובחשיבותם –

* **צדק חברתי** – בעזרת הנתונים נוכל להדגיש פערים באופן בו קבוצות דמוגרפיות שונות מושפעות מאלימות משטרתית. על ידי הבנה וזיהוי של פערים אלו נוכל להבין בצורה טובה יותר בעייתיות מערכתית במשטרה, ולפעול להבטיח שכל הקהילות יקבלו יחס שוויוני.
* **תיקון מדיניות משטרתית** – קובעי המדיניות יכולים להשתמש בתובנות הפרויקט בכדי ליצור חוקים ותקנות יעילים יותר שמטרתם להפחית ולדלל מקרים של אלימות משטרתית.
* **קהילתיות** – שיפור היחסים בין רשויות אכיפת החוק והקהילות שהם משרתות הוא חיוני לביטחון הציבור ואמונם. הבנת הסיבות והמגמות של רציחות המשטרה יכולה לעזור לגשר על פערים ולטפח תקשורת ושיתוף פעולה טובים יותר.
* **מניעה** – על ידי בחינת הגורמים הסוציו אקונומיים, ניתוח המידע יכול לזהות אזורים בעלי סיכון גבוה יותר להרג משטרתי, מידע זה יכול להנחות הקצאת משאבים לאמצעי מניעה כגון תוכניות תמיכה קהילתיות ויוזמות לפיתוח כלכלי.

על ידי ניתוח מידע זה נוכל לשפר את מצב הקיים ולענות של השאלה המרכזית שלנו שהיא – מציאת הדפוסים והמגמות הכוללות של הרציחות המשטרתיות ברחבי ארה"ב ב-2015, וכיצד מגמות אלו מתואמות עם גורמים דמוגרפיים, גיאוגרפיים וסוציו אקונומיים שונים.

הניתוח שלנו מבוסס על מערך נתונים מקיף של רציחות משטרתיות ב-2015, הכולל מידע מפורט על כל אירוע. מערך הנתונים כולל 34 עמודות שונות ביניהם גיל, מגדר, גזע הקורבנות, תאריך ומיקום כל אחד האירועים וכו'. המידע בנוסף לכך כולל אינדיקטורים סוציו אקונומיים של האזורים בהם התרחש האירוע כגון רמות הכנסה, שיעורי עוני ושיעורי אבטלה. מערך נתונים עשיר זה מאפשר לנו לבצע ניתוח רב גוני כדי לחשוף דפוסים ומגמות משמעותיות.

נבצע טבלה בכדי לתאר את המידע, כאשר המזהה של הטבלה ( מפתח ראשי) הינו שם הנרצח (Name). הטבלה תכיל את שם העמודה, הקרדינליות שלה (כמות הערכים הייחודים), תיאור של אותה עמודה והמשמעות שלה עבור הפרויקט ועבור הויזואליזציות שנבצע:

|  |  |  |  |
| --- | --- | --- | --- |
| **עמודה** | **קרדינליות** | **תיאור** | **משמעות** |
| Name | 467 | שם האדם שנהרג על ידי המשטרה | עוזר בזיהוי המקרים הבודדים |
| age | 60 | גיל האדם שנהרג על ידי המשטרה | שימושי לניתוח דמוגרפי |
| gender | 2 | מגדר האדם שנהרג על ידי המשטרה | מאפשר ניתוח לפי מגדר |
| raceethnicity | 6 | גזע/אתניות של האדם שנהרג על ידי המשטרה | קריטי להבנת הפערים הגזעיים |
| |  | | --- | |  |   month, day, year | 12 (month), 31 (day), 1 (year) | תאריך האירוע | מאפשר ניתוח סדרות זמן לזיהוי מגמות לאורך הזמן |
| streetaddress, city, state | 467 (add),  247 (city),  34 (state) | פרטי מיקום האירוע | חשוב לניתוח גיאוגרפי ומיפוי האירועים |
| latitude, longitude | 466 | קואורדינטות גיאוגרפיות של מיקום האירוע | עוזר בזיהוי דפוסים הקשורים לסוכנויות ספציפיות |
| lawenforcementagency | 117 | סוכנות אכיפת החוק המעורבת באירוע | מספק תובנה לגבי נסיבות כל אירוע |
| cause | 10 | סיבת המוות | מספק תובנה לגבי נסיבות כל אירוע |
| armed | 5 | האם האדם היה חמוש | חשוב להבנת ההקשר של האירוע |
| pop | 463 | אוכלוסיית האזור | מספק הקשר במונחי צפיפות אוכלוסייה |
| share\_white, share\_black, share\_hispanic | 463 | אחוז האוכלוסייה הלבנה, השחורה וההיספנית באזור | שימושי להקשר דמוגרפי ולניתוח |
| p\_income | 463 | הכנסה לנפש באזור | גורם סוציו-אקונומי לניתוח |
| h\_income | 463 | הכנסה ביתית באזור | גורם סוציו-אקונומי נוסף לניתוח |
| county\_income | 43 | הכנסה ממוצעת במחוז | מספק הקשר ברמת המחוז |
| comp\_income | 463 | השוואת הכנסה ביתית לממוצע המחוז | עוזר לזהות אזורים עם פערים כלכליים |
| county\_bucket, nat\_bucket | 5 | קטגוריות הכנסה במחוז וברמה הלאומית | שימושי לסיווג סוציו-אקונומי |
| pov | 461 | שיעור העוני באזור | אינדיקטור סוציו-אקונומי חשוב |
| urate | 460 | שיעור האבטלה באזור | מספק הקשר כלכלי |
| college | 463 | אחוז האוכלוסייה עם השכלה גבוהה באזור | אינדיקטור להשכלה |

**כעת נמפה את המידע לפי הטיפולוגיה האבסטרקטית של Munzner-**

Dataset Type – מערך הנתונים הוא טבלה שבה כל שורה מייצגת אירוע מסוים (הרג על ידי שוטרים), והעמודות מייצגות תכונות של אירועים אלו.

|  |  |
| --- | --- |
| **Data Types** | **Columns** |
| Categorical | name, gender, raceethnicity, cause, armed, lawenforcementagency |
| Quantitative | age, latitude, longitude, pop, share\_white, share\_black, share\_hispanic, p\_income, h\_income, county\_income, comp\_income, pov, urate, college |
| Temporal | month, day, year |
| Geographical | streetaddress, city, state, latitude, longitude |

|  |  |
| --- | --- |
| **Attributes** | **Columns** |
| Nominal | name, gender, raceethnicity, cause, armed, lawenforcementagency, city, state |
| Ordinal | county\_bucket, nat\_bucket |
| Quantitative | age, latitude, longitude, pop, share\_white, share\_black, share\_hispanic, p\_income, h\_income, county\_income, comp\_income, pov, urate, college |
| Temporal | month, day, year |
| Geographical | streetaddress, latitude, longitude |

נתאר את מטלות המשתמש בעזרת כמה שאלות ונמפה אותם למונחים אבסטרקטים בהתאם לטיפולוגיה של Munzer –

|  |  |
| --- | --- |
| **מטלת משתמש** | **מיפוי** |
| מהי התפלגות הרג המשטרה לפי גזע ומוצא אתני עם פילוח מגדרי? | {action = Compare, target =Distribution} |
| האם יש מדינות או ערים ספציפיות עם שיעור גבוה יותר של הרג משטרתי? | {action = Identify, target =Spatial region} |
| כיצד מתקשר המצב הסוציו אקונומי של אזור מחייה לתדירות הרציחות המשטרתיות? | {action = Correlate, target =Relationship} |
| מהן המגמות לאורך זמן ברציחות משטרתיות? | {action =Summarize, target =Trend} |

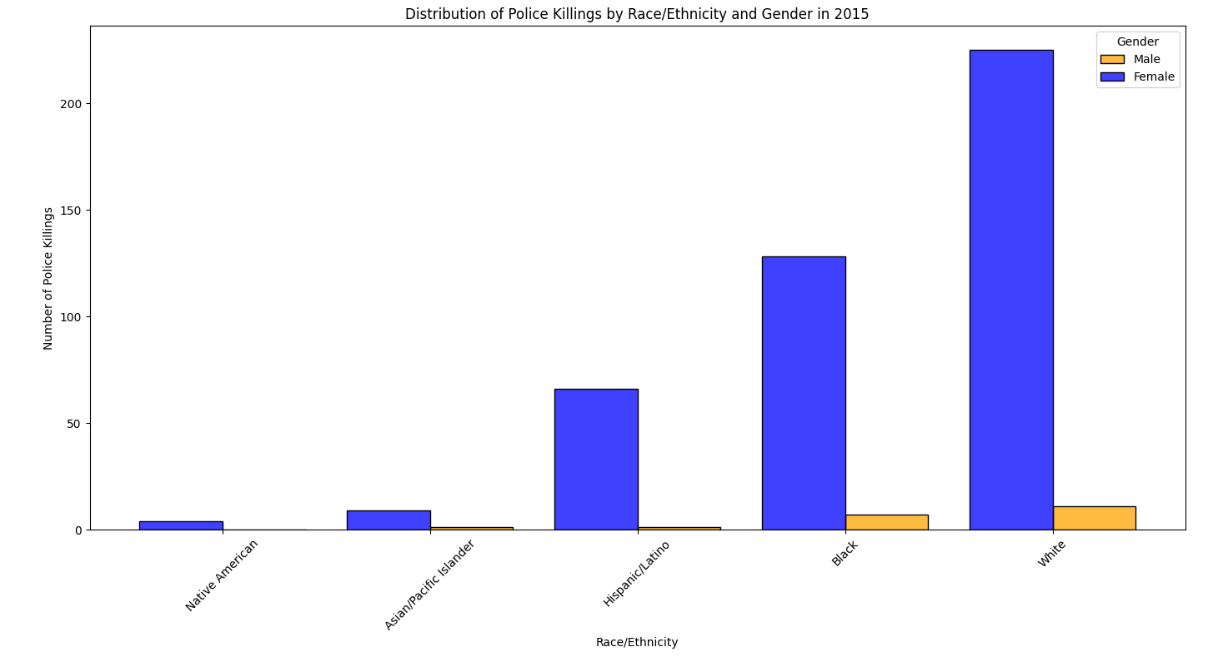
עיצובים חלופיים-

בחלק זה של הפרויקט אנו נבצע סקירה של מספר חלופות עיצוביות לויזואליזציה של הנתונים אודות הרג המשטרתי בארצות הברית בשנת 2015. המטרה המרכזית היא לבחון מספר גישות שונות להצגת המידע, על מנת לזהות את החלופות היעילות ביותר בהעברת המסר וניתוח הנתונים. בתהליך זה נציג שתי חלופות ויזואליות, נבצע השוואה בין היתרונות והחסרונות של כל חלופה ונבחר את החלופה המתאימה ביותר על סמך קריטריונים מוגדרים מראש.

שאלה ראשונה שנרצה לענות עליה היא - מהי התפלגות הרג המשטרה לפי גזע ומוצא אתני עם פילוח מגדרי? ועליה נענה בעזרת שני חלופות ויזואליזציה שונות.

חלופה ראשונה: גרף עמודות.

סקיצה של החלופה-



**marks:**

עמודות: כל עמודה מייצגת את מספר ההרוגים לכל קטגוריה של גזע/מוצא אתני ומגדר.

**channels:**

גובה העמודה: מייצג את מספר ההרוגים לכל קטגוריה של גזע/מוצא אתני ומגדר.

צבע העמודה: מייצג את המגדר (כחול לנשים, כתום לגברים).

מיקום אופקי: מייצג את הגזע/מוצא אתני.

**user task:**

השוואה: המשתמש צריך להשוות את מספר ההרוגים בין הקבוצות השונות של גזע/מוצא אתני ומגדר.

זיהוי דפוסים: המשתמש צריך לזהות דפוסים בתדירות ההרוגים לפי גזע/מוצא אתני ומגדר.

סידור: המשתמש יכול לראות את הסידור של הקבוצות מהקטנה לגדולה בהתבסס על מספר ההרוגים הכולל.

**חלופה שנייה:** גרף עוגה.

סקיצה של החלופה -



**marks:**

פלחים בעוגה: כל פלח בעוגה מייצג את אחוז ההרוגים לפי גזע/מוצא אתני ומגדר.

**channels:**

גודל הפלח: מייצג את אחוז ההרוגים מכלל הקבוצה של גזע/מוצא אתני עבור כל מגדר.

צבע הפלח: מייצג את המגדר (כחול לנשים, כתום לגברים).

תוויות: מציגות את האחוזים עבור כל פלח.

**user task:**

זיהוי דפוסים: המשתמש צריך לזהות דפוסים ביחס בין המגדרים השונים בתוך כל קבוצת גזע/מוצא אתני.

השוואה: המשתמש צריך להשוות את האחוזים של ההרוגים בין המגדרים השונים בתוך כל קבוצת גזע/מוצא אתני.

הבנה כללית: המשתמש מקבל תמונה כללית על היחס בין המגדרים בכל קבוצת גזע/מוצא אתני ובין הקבוצות השונות.

**השוואת הויזואליזציות -**

|  |  |  |
| --- | --- | --- |
| **קריטריון** | **חלופה ראשונה - גרף עוגה** | **חלופה שנייה - גרף עמודות** |
| אפקטיביות | אפקטיביות השטח - העוגה מאפשרת להמחיש את החלוקה היחסית בין המינים עבור כל גזע/מוצא אתני מה שעוזר להבין את האחוזים בקלות.  אפקטיביות הצבע: צבעים שונים מסייעים להבחין בקלות בין מין הנרצחים בכל קטגוריה. | אפקטיביות הגובה- קל להשוות בין כמות ההרוגים של כל גזע/מוצא אתני באמצעות גובה העמודות.  אפקטיביות הצבע - הצבעים מאפשרים להבחין בקלות בין המינים בתוך כל קטגוריה. |
| אקספרסיביות | העוגה אקספרסיבית מספיק כדי להראות את ההבדלים היחסים בין המינים בתוך כל גזע/מוצא אתני. עם זאת, היא פחות אקספרסיבית כשמדובר בהשוואות בין קבוצות שונות. | העמודות אקספרסיביות מאוד להצגת הבדלים במספר ההרוגים בין גזעים/מוצאים אתניים שונים ובין המינים. |
| יתרונות | קל להבין את ההבדלים באחוזים בין המינים עבור כל גזע/מוצא אתני.  מתאים להצגת חלוקה יחסית. | קל להשוות בין כמות ההרוגים של כל גזע/מוצא אתני.  קל לראות את ההבדלים בין המינים. |
| חסרונות | קשה להשוות בין סקטורים קטנים במיוחד.  פחות מתאים להשוואה ישירה בין קבוצות. | קל להשוות בין כמות ההרוגים של כל גזע/מוצא אתני.  קל לראות את ההבדלים בין המינים. |

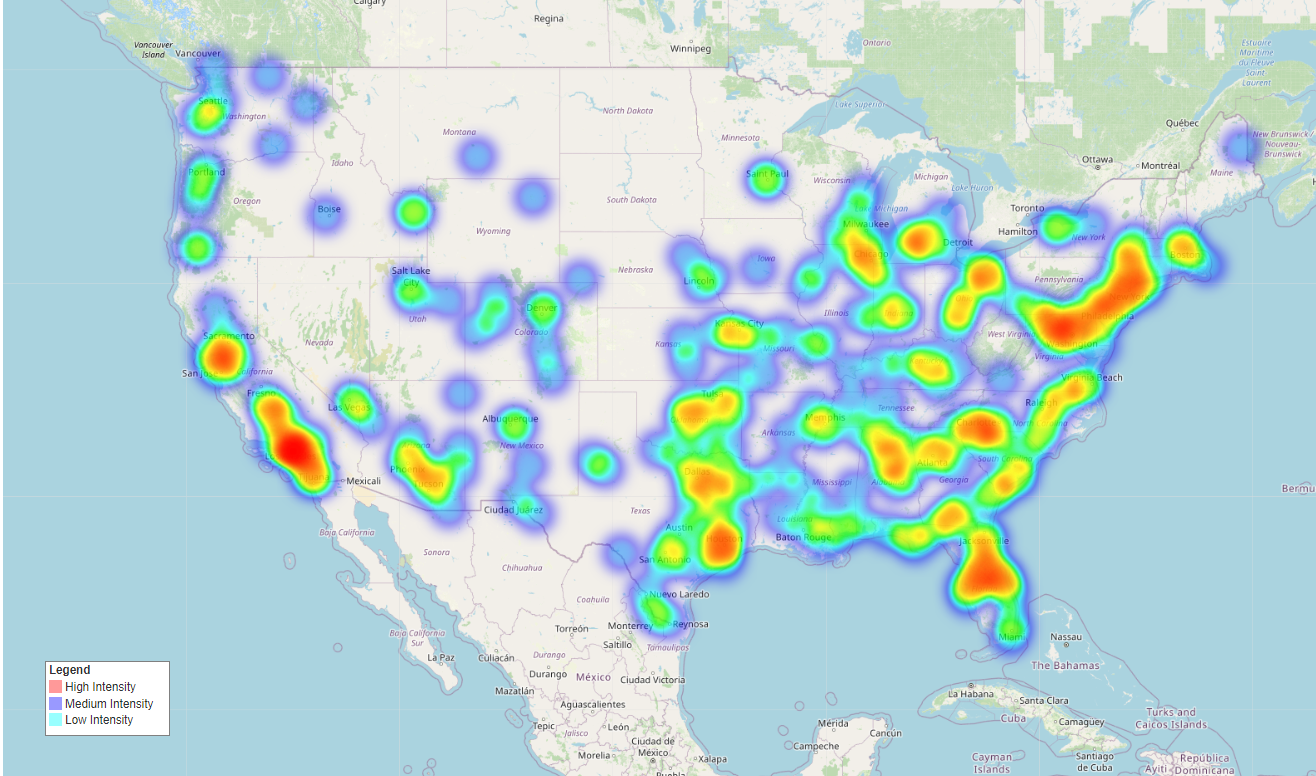
איזה ויזואליזציה בחרנו ולמה?

אנו נבחר את אופציה א' בכדי לייצג שאלה זאת, מכיוון שהיא מספקת בהירות וקריאות ברורות. באמצעות גרף עמודות, ניתן לראות בצורה מיידית את מספר ההריגות בכל קטגוריה של גזע/מוצא אתני ומגדר. גובה העמודות מאפשר השוואה ברורה בין הקבוצות השונות והצבעים השונים עבור זכרים ונקבות מקלים על ההבחנה ביניהם. בנוסף, העמודות מבטאות בצורה אפקטיבית את המספרים האבסולוטיים, מה שמדגיש את ההבדלים והפערים בין הקבוצות. הבחירה בגרף עמודות מסייעת להבין את ההבדלים בממדים השונים ובכך מעודדת ניתוח מעמיק יותר של הנתונים והבנה טובה יותר של התופעה המדוברת.

שאלה 2: האם יש מדינות או ערים ספציפיות עם שיעור גבוה יותר של הרג משטרתי?

חלופה ראשונה: מפה אינטראקטיבית.

סקיצה של החלופה -



**marks:**

נקודות : כל נקודה על המפה מייצגת מיקום של הרג משטרתי.

**channels:**

צבע:הצבעים במפת החום מצביעים על עוצמת ההריגות באזורים שונים.

צבעים אדומים מייצגים אזורים עם שיעור גבוה יותר של הרג משטרתי.

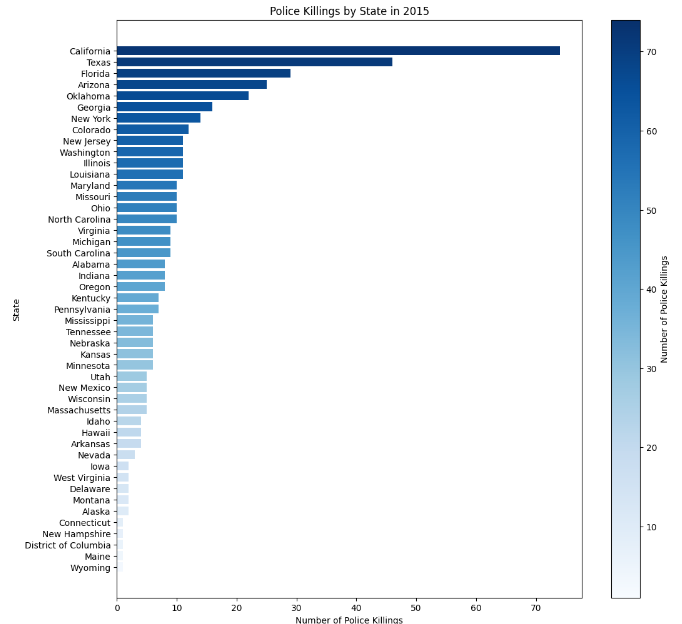
צבעים כחולים מייצגים אזורים עם שיעור נמוך יותר של הרג משטרתי.

**user task:**

זיהוי : המשתמש יכול לזהות אזורים עם שיעור גבוה של הרג משטרתי על פי הצבעים החמים (אדומים) במפת החום, המשתמש יכול לזהות אזורים עם שיעור נמוך של הרג משטרתי על פי הצבעים הקרים (כחולים) במפת החום.

השוואה: המשתמש יכול להשוות בין אזורים שונים בארה"ב ולראות היכן יש עוצמות גבוהות יותר של הרג משטרתי.

מיקום: המשתמש יכול למקם את האזורים הגיאוגרפיים הספציפיים בהם יש שיעור גבוה יותר של הרג משטרתי על המפה ולבחון את הפיזור הגיאוגרפי שלהם.

חלופה שנייה: גרף עמודות

**marks:**

עמודות: המייצגות כל מדינה.

טקסט**:** עבור שמות המדינות.

**channels:**

אורך :המייצג את מספר ההריגות המשטרתיות.

רוויית צבע: של העמודות מציינת את עוצמת ההרג המשטרתי.

מיקום :המייצג מדינות שונות המסודרות מהכי מעט להכי הרבה הריגות משטרתיות.

**user task:**

זיהוי : לזהות את המדינות עם המספר הגבוה והנמוך ביותר של הריגות משטרתיות ב-2015.

השוואה: להשוות את מספר ההריגות המשטרתיות בין המדינות השונות כדי לראות אילו מדינות יש בהן יותר או פחות הריגות.

פירוש: לנתח ולפרש את התפלגות ועוצמת ההריגות המשטרתיות לפי מדינה בשנת 2015 כדי להבין את ההבדלים האזוריים.

**השוואת הויזואליזציות-**

|  |  |  |
| --- | --- | --- |
| **קריטריון** | **חלופה ראשונה - מפה אינטראקטיבית** | **חלופה שנייה - גרף עמודות** |
| אפקטיביות | מפה חמה היא כלי יעיל לזיהוי דפוסים מרחביים במידע.  מאפשרת זיהוי מהיר של אזורים עם ריכוז גבוה של מקרי הרג. | גרף עמודות מאפשר השוואה ישירה וברורה בין מדינות.  קל לזהות את המדינות עם המספר הגבוה ביותר של מקרי הרג. |
| אקספרסיביות | המפה החמה אקספרסיבית בכך שהיא מציגה את ההתפלגות המרחבית של מקרי ההרג.  השימוש בצבעים שונים מאפשר להבחין בין אזורים עם ריכוזים שונים של מקרי הרג בצורה ברורה. | הגרף אקספרסיבי בכך שהוא מציג את מספר מקרי ההרג בצורה ברורה וישירה.  השימוש בצבעים מגביר את היכולת לזהות הבדלים במספר מקרי ההרג בין המדינות. |
| יתרונות | מאפשרת זיהוי דפוסים מרחביים בצורה ברורה ומיידית.  מאפשרת הצגת כמות רבה של מידע במבט אחד. | קל לקרוא ולהבין.  מאפשר השוואה ישירה בין מדינות. |
| חסרונות | דורשת מיקום גיאוגרפי מדויק של כל אירוע.  קשה לראות את הפרטים המדויקים באזורים קטנים.  לא מאפשר לראות מספר מדויק של נרצחים בכל מדינה. | לא מציג את הפריסה הגיאוגרפית של מקרי ההרג.  פחות מתאים להצגת דפוסים מרחביים. |

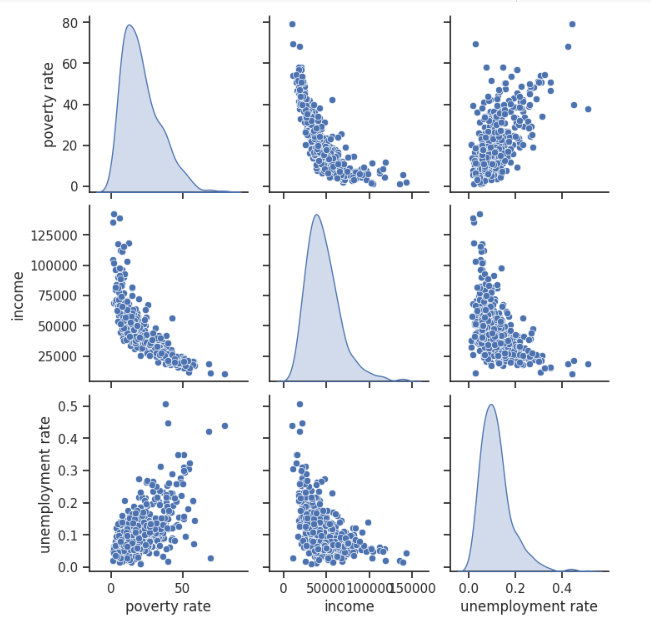
איזה ויזואליזציה בחרנו ולמה?

אנו נבחר את אופציה א' בכדי לייצג שאלה זאת, מכיוון שהמפה האינטראקטיבית מאפשרת זיהוי מיידי של אזורים עם שיעור גבוה של מקרי הרג משטרתי ברחבי ארצות הברית. באמצעות הצבעים החמים (אדום) ניתן לראות במהירות ובבהירות את הריכוזים הגבוהים ביותר של מקרי הרג, בעוד שהצבעים הקרים (כחול) מראים את האזורים עם שיעור נמוך יותר. השימוש במפה זו מאפשר גם הבנה גיאוגרפית מדויקת של פריסת מקרי ההרג ומקל על זיהוי הערים והמדינות הספציפיות שבהן נדרשת התערבות ושיפור. בנוסף, המפה מציגה את הנתונים בצורה אסתטית ואינטראקטיבית, מה שמגביר את האפקטיביות של ההצגה ואת היכולת לזהות מגמות ודפוסים ברמה הארצית.

שאלה 3: כיצד מתקשר המצב הסוציו אקונומי של אזור מחייה לתדירות הרציחות המשטרתיות?

חלופה ראשונה: scatter plot matrix

סקיצה של החלופה -



**marks:**

נקודות: מייצגות את הערכים של כל המשתנים במערכת צירים קרטזית.

עקומות צפיפות: מייצגות את התפלגות הערכים של כל משתנה בנפרד

**channels:**

מיקום: מציג את הקורלציה בין המשתנים.

סדר: מציג את השינויים ברמות השונות של המשתנים.

צפיפות נקודות ועקומות: מציגה את הקורלציה וההצטברות של הערכים.

**user task:**

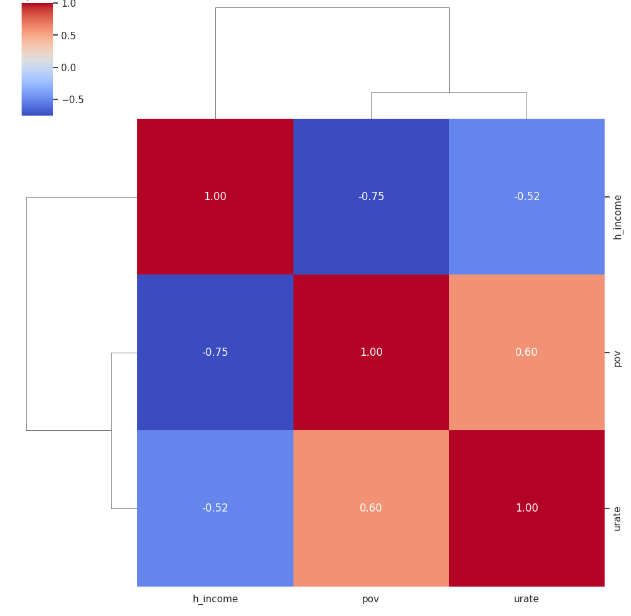
זיהוי: לזהות את התפלגות ערכי המשתנים השונים.

השוואה: להשוות בין ערכי המשתנים השונים ולראות את הקורלציה ביניהם.

פירוש: לנתח את היחסים בין המשתנים כדי להבין כיצד הם משפיעים זה על זה.

חלופה שנייה: heat map

סקיצה של החלופה-



**marks:**

ריבועים: מייצגים את ערכי הקורלציה בין המשתנים.

**channels:**

צבע: מציג את עוצמת הקורלציה בין המשתנים, עם סקאלה מ-1 (אדום כהה) עד -1 (כחול כהה).

טקסט: ערכי הקורלציה מוצגים על גבי הריבועים.

מיקום: מייצג את הקשר בין המשתנים השונים בשתי צירי המטריצה.

**user task:**

זיהוי: לזהות את רמת הקורלציה בין המשתנים השונים.

השוואה: להשוות בין ערכי הקורלציה ולראות את הקשרים החזקים והחלשים ביותר.

פירוש: לנתח את היחסים בין המשתנים כדי להבין את הקשר ביניהם.

|  |  |  |
| --- | --- | --- |
| **קריטריון** | **חלופה ראשונה - scatter plot matrix** | **חלופה שנייה - heat map** |
| אפקטיביות | scatter plot מאפשר השוואה ישירה וברורה בין המשתנים הסוציו-אקונומיים.  מאפשר זיהוי מהיר של הקורלציות בין המשתנים הסוציו-אקונומיים השונים כפונקציה של מספר מקרי הרציחות. | heat map היא כלי יעיל לזיהוי דפוסים מרחביים במידע.  מאפשרת זיהוי מהיר של קורלציות בין המשתנים הסוציו-אקונומיים. |
| אקספרסיביות | הגרף אקספרסיבי בכך שהוא מציג את הקורלציות בצורה ברורה וישירה.  השימוש בנקודות וצבעים מגביר את היכולת לזהות הבדלים הצטברות בין המשתנים. | heat map אקספרסיבית בכך שהיא מציגה את עוצמת הקורלציה בין המשתנים.  השימוש בצבעים שונים מאפשר להבחין בין קורלציות חיוביות ושליליות בצורה ברורה. |
| יתרונות | מאפשר השוואה ישירה וברורה בין משתנים שונים.  מציג את הקורלציות בצורה גרפית ומובנת. | מאפשרת זיהוי דפוסים בצורה ברורה ומיידית.  מציגה את עוצמת הקורלציה בצורה ברורה וישירה. |
| חסרונות | עשוי להיות מורכב מדי לצפייה כשיש הרבה משתנים.  פחות מתאים להצגת דפוסים מרחביים. | דורשת פרשנות צבעים כדי להבין את הקורלציות.  קשה לראות את הפרטים המדויקים כשיש הרבה משתנים עם קורלציות נמוכות.  לא מאפשר לייצג קשר בין המשתנים השונים לבין מספר מקרי הרציחות. |

איזה ויזואליזציה בחרנו ולמה?

אנו נבחר את אופציה א' בכדי לייצג שאלה זאת, מכיוון ש-Scatter Plot Matrix מאפשרת לראות בבירור את הקורלציות בין המשתנים השונים וכיצד הם משפיעים זה על זה. השימוש בגרפים אלו מציג את הקשרים בין כל זוג משתנים בצורה ברורה ומובנת, ומסייע בזיהוי מגמות ודפוסים משמעותיים. בנוסף, ניתן לראות את ההתפלגויות של הערכים עבור כל משתנה בנפרד, מה שמוסיף רובד נוסף של הבנה לגבי המבנה של הנתונים. הבחירה ב-Scatter Plot Matrix תורמת להבנה מעמיקה יותר של הנתונים ומאפשרת זיהוי מהיר של הקשרים החזקים והחלשים ביותר בין המשתנים.

חשוב לציין כי השאלה הרביעית היא השאלת הdashboard שלנו, ושלושת השאלות עליהם ביצענו ויזואליזציה מסכמות שאלה זאת ועונות עליה בדרך כזאת או אחרת.

הסבר על העיצוב שנבחר ויישומו -

### מפה אינטראקטיבית

#### שלבי עיבוד מקדימים

* ניקוי נתונים: קובץ הנתונים נטען וכל השורות עם הערכים החסרים או לא מספריים של קווי רוחב ואורך הוסרו. דבר זה מבטיח לנו שהמפה החמה תייצג באופן מדויק את הפיזור הגיאוגרפי של מקרי ההרג על ידי המשטרה.
* אימות נתונים: ערכי קווי הרוחב והאורך אומתו כדי לוודא שהם נמצאים בטווחים התקינים (מ-90- עד 90 עבור קו רוחב ומ-180- עד 180 עבור קו אורך).

#### בחירות ויזואליות

* מרכוז המפה: המפה ממורכזת מעל ארצות הברית עם רמת זום מתאימה כדי לספק מבט ברור על נקודות הנתונים.
* מפה חמה: שכבת מפה חמה נוספה כדי להמחיש את ריכוז מקרי ההרג של המשטרה גיאוגרפית. ריכוזים גבוהים מוצגים עם צבעים אינטנסיביים יותר.
* מקרא: מקרא נוסף למפה כדי לעזור למשתמשים לפרש את האינטנסיביות של צבעי המפה החמה, המציינת אזורים בעלי אינטנסיביות נמוכה, בינונית וגבוהה.

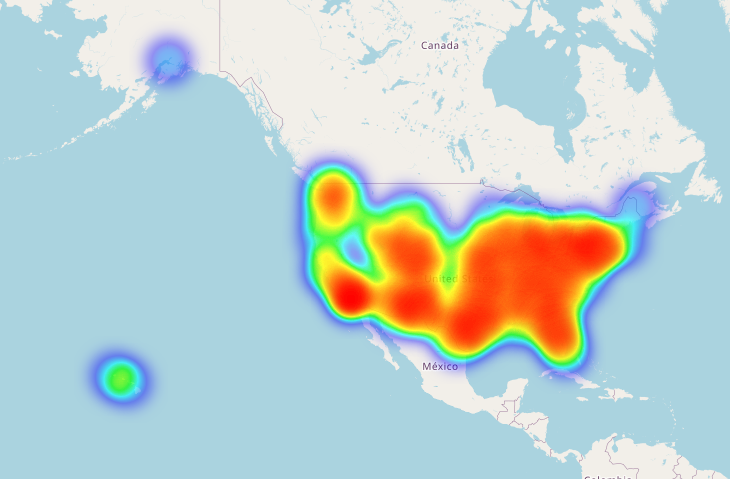
#### אינטראקטיביות

* אלמנטים אינטראקטיביים: המפה מאפשרת למשתמשים לבצע זום וגרירה, מספקת דרך דינמית לחקור את הנתונים במרחב.
* הטמעת Iframe: המפה מוטמעת בתוך Iframe באפליקציית Dash, המאפשרת למשתמשים לקיים אינטראקציה איתה ישירות בדפדפן.

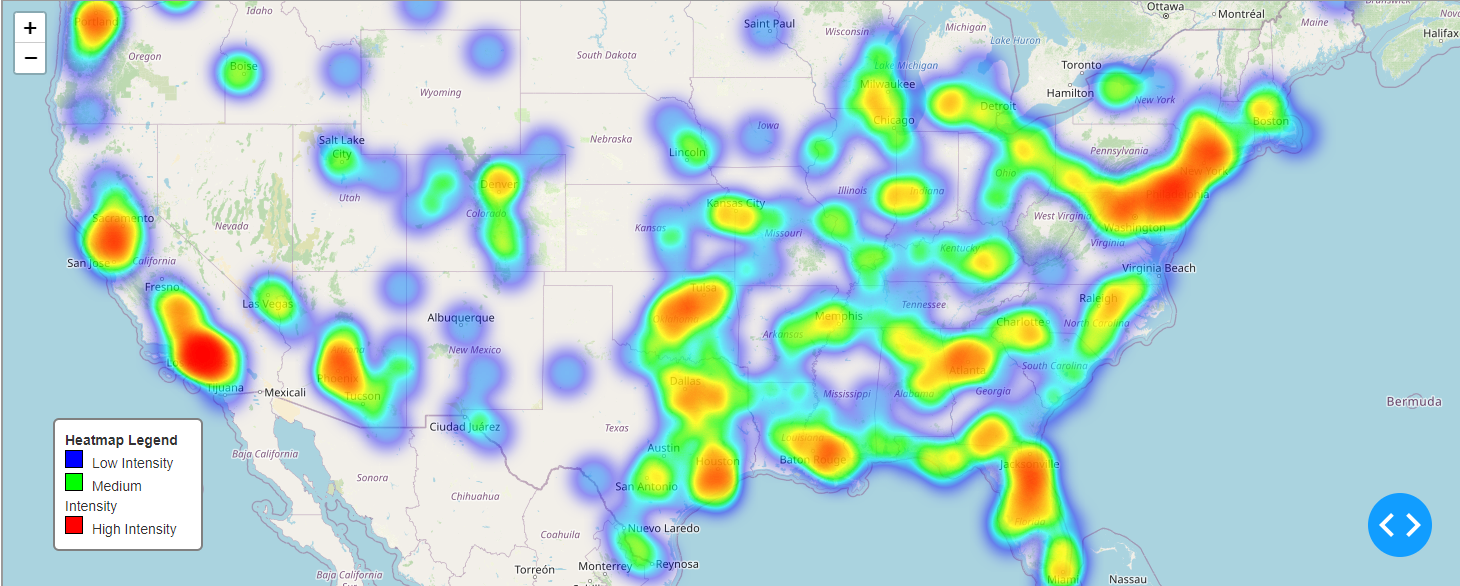
#### ספריות קוד בשימוש

* Dash: ליצירת אפליקציה אינטראקטיבית.
* Folium: ליצירת המפה והוספת שכבת המפה החמה.
* Pandas: לטיפול בעיבוד הנתונים.

הוראות שימוש במפה

****

ניתן להבחין במבט ראשוני שהמפה מסתכלת על כלל אזורי ארצות הברית, אך בעזרת הזזת עכבר פשוטה אפשר לשנות אזורים בארצות הברית ולהבחין בצורה טובה את ההבדלים באזורים מבחינת רציחות בארצות הברית.



בעזרת גלגלת העכבר אפשר להתקרב ולראות את מפת החום משתנה לפי האזורים, לפי המקרא להבין באלו אזורים הרציחות בארצות הברית היו בכמות יותר גדולה ופחות גדולה.

**גרף עמודות**

**שלבי עיבוד מקדימים**

* ניקוי נתונים: הקובץ נטען וטורים כמו גיל ו-"race/ethnicity" נוקו מערכים חסרים או לא מספריים. כל קטגוריות ה-"Unknown" מכיוון שלא מספקות שום מידע נוסף.
* קיבוץ גילאים: הגילאים חולקו לקבוצות גיל כדי לאפשר ניתוח מקיף יותר והבנה ברורה יותר של הפילוג.
* עמודת 'armed' בינארית: עמודה בינארית נוצרה עבור נתוני ה-"armed" לצורך ניתוח פשוט וברור יותר וקריאה ברורה יותר של הגרף.

**בחירות ויזואליות**

* גרף עמודות קבוצתי: גרף עמודות נבחר כדי להמחיש את ההתפלגות לפי גזע/אתניות ומשתנים נוספים כמו גיל, מגדר, חודש, וחמוש/לא חמוש.
* סדר קטגוריות: סדר הקטגוריות נקבע מראש כדי להציג את הנתונים בסדר לוגי וברור.
* סולם צבעים: שימוש בסולם צבעים מוגדר מראש להדגשת הקבוצות השונות בצורה ויזואלית והבנה ברורה יותר שלהם.

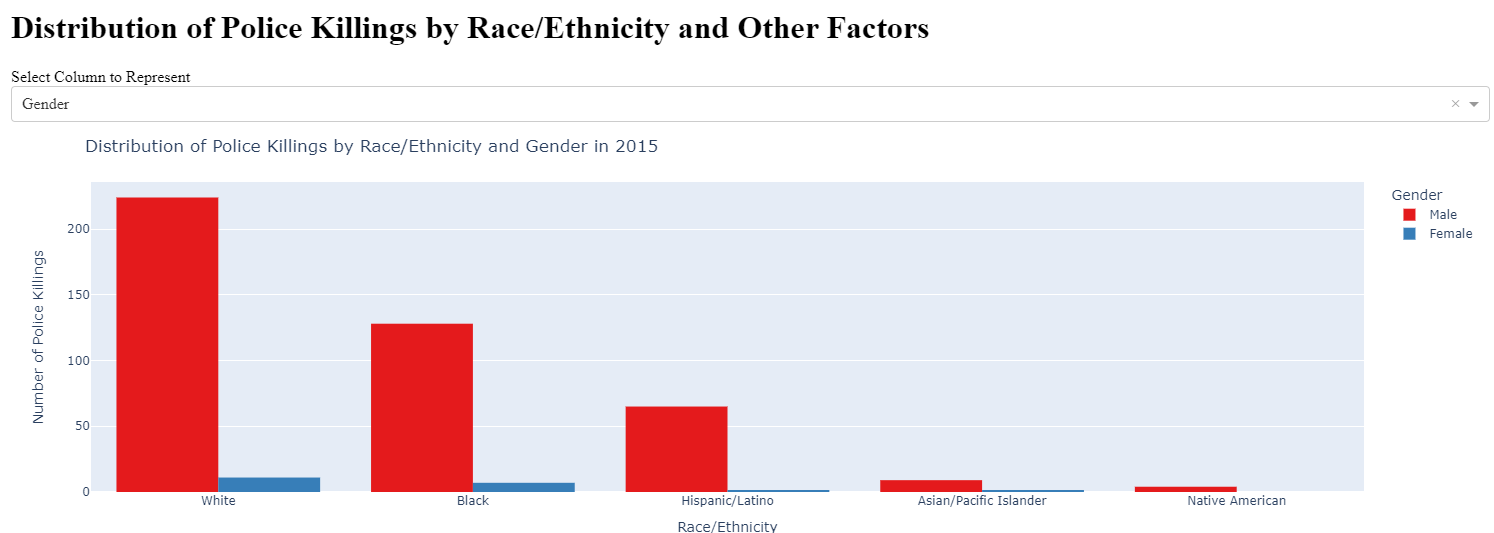
**אינטראקטיביות**

* Dropdown אינטראקטיבי: משתמשים יכולים לבחור משתנים שונים לייצוג בגרף באמצעות תיבת Dropdown אינטראקטיבית.
* גרף דינמי: הגרף מתעדכן באופן דינמי בהתאם לבחירת המשתמש, מה שמאפשר ניתוח גמיש והקפי יותר של הנתונים.

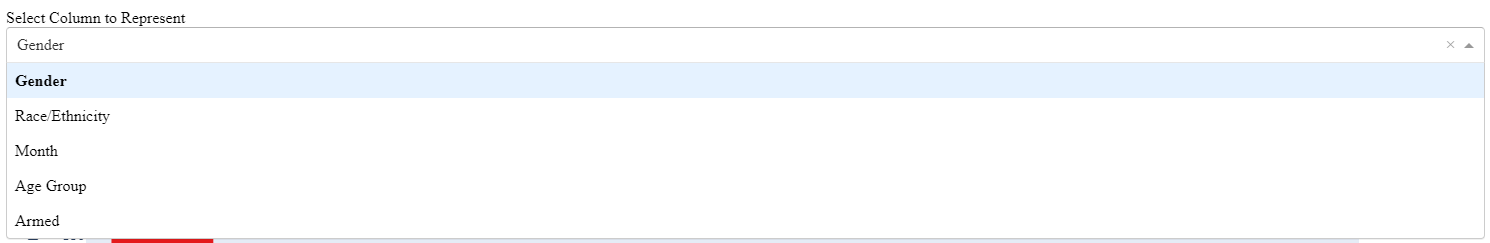
**ספריות קוד בשימוש**

* Dash : ליצירת אפליקציה אינטראקטיבית.
* Plotly Express: ליצירת גרפי עמודות אינטראקטיביים.
* :Pandas לטיפול נוח יותר בעיבוד הנתונים.

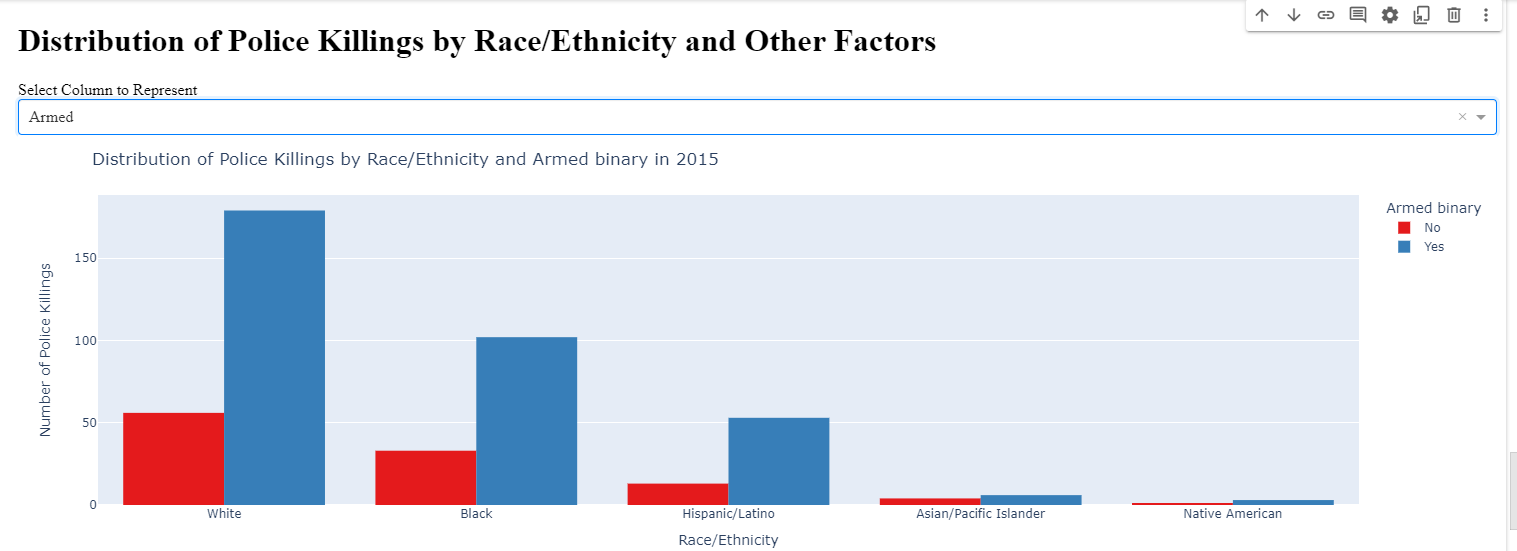
הוראות שימוש בגרף



באופן ברור מאוד ניתן לבחור בעזרת לחיצה פשוטה באיזה מן עמודה נרצה לפלג את הנתונים.



בעזרת לחיצה ברשימה זאת נוכל לבחור לפי איזה משתנה נרצה לפלג, דוגמא נוספת –



**מטריצה של משתנים סוציו-אקונומיים**

**שלבי עיבוד מקדימים**

* ניקוי נתונים: הנתונים הוסבו לערכים מספריים על מנת להבטיח ניתוח נכון.
* בחירת משתנים :המשתנים הסוציו-אקונומיים הרלוונטיים נבחרו להכללה במטריצה.

**בחירות ויזואליות**

* מטריצת :Pair Plot מטריצה של גרפים עם התפלגות בצד האלכסון וגרפי פיזור בין המשתנים, כדי להמחיש את הקשרים האפשריים בין המשתנים הסוציו-אקונומיים השונים תוך הבנת התפלגות הנתונים עצמם.
* שמות אינפורמטיביים: שמות העמודות הוסבו לשמות יותר אינפורמטיביים ומובנים.

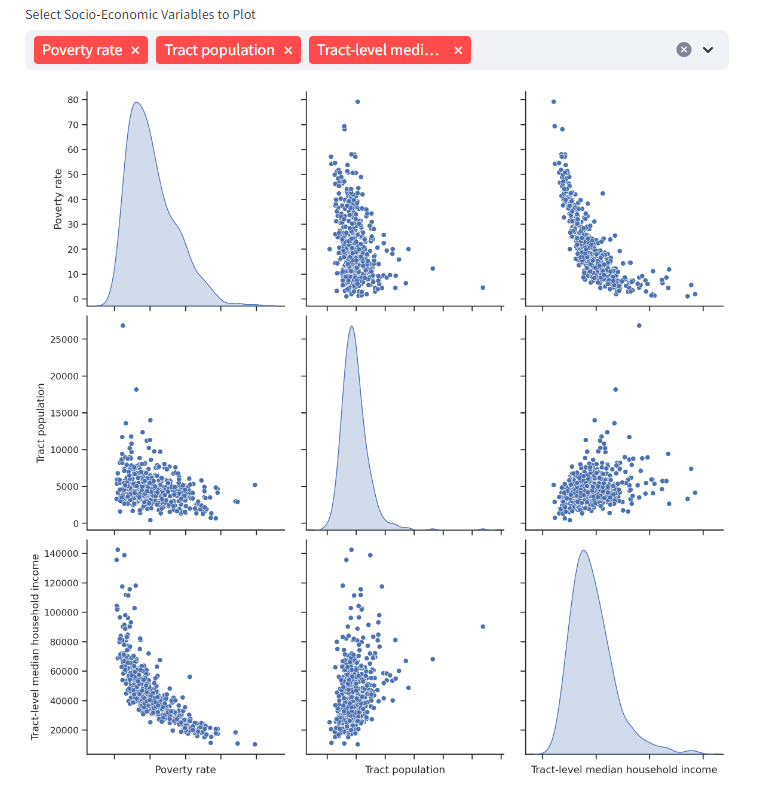
**אינטראקטיביות**

* Dropdown אינטראקטיבי: משתמשים יכולים לבחור אילו משתנים לכלול במטריצה באמצעות תיבת Dropdown אינטראקטיבית.
* עדכון גרף דינמי: המטריצה מתעדכנת באופן דינמי בהתאם לבחירת המשתמש, מה שמאפשר ניתוח מותאם אישית של הנתונים.

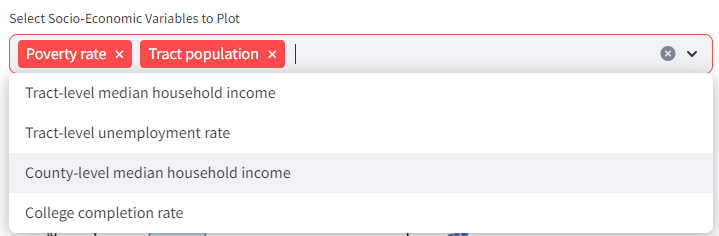
**ספריות קוד בשימוש**

* Dash : ליצירת אפליקציה אינטראקטיבית.
* Seaborn : ליצירת מטריצת Pair plot .
* Pandas: לטיפול בעיבוד הנתונים.

הוראות שימוש בגרף



אלו הערכים שמקבלים כברירת מחדל בגרף, ניתן להוריד משתנים בעזרת הלחצן שמופיע ליד כל שם, וניתן להוסיף בעזרת הרשומה הנגללת. מומלץ להציג בין 2 ל-3 משתנים בכדי שהגרף יהיה ברור ככל הניתן.



האופציות שניתנות לבחירה והוספה אל הגרף מצורפות לעיל.

דוגמה נוספת של הגרף –

