

## תאונות הדרכים בישראל

## קישור לDashboard: לחצו כאן

שם הקורס: ויזואליזציה של מידע.

תאריך ההגשה: 20/07/2024

שמות המגשים:

**שם:** דור מאיר ; **ת"ז:** 313254724

**שם:** סליון נעמי איבשין, **ת"ז:** 322769175



## מוטיבציה קטנה:

תאונות דרכים הן אחת מהבעיות הבולטות והקשות ביותר בתחום הבטיחות הציבורית, והן גורמות לאובדן חיי אדם, פציעות קשות ונזק כלכלי רב. על ידי חקר הנושא והבנת הגורמים השונים לתאונות, ניתן לפתח פתרונות ושיטות לשיפור הבטיחות בדרכים ולהפחתת מספר התאונות. בנוסף, נושא זה משלב בין מדעי החברה, טכנולוגיה והנדסה, ומאפשר לפתח גישה רב-תחומית להתמודדות עם הבעיה. יש חשיבות רבה למאמץ הלאומי לשיפור הבטיחות בדרכים ולהצלחת דרכים. נושא תאונות הדרכים הינו בעל ערך מוסרי וחברתי גבוה.

## חלק 2: ניתוח נתונים ומטלות:

### 1. מבוא

נושא הויזואליזציה: תאונות דרכים במדינת ישראל

השאלה המרכזית עליה אנו מבקשים לענות: כיצד מאפיינים סביבתיים שונים השפיעו על תאונות הדרכים בישראל? נרצה לחקור את הנתונים כך שנוכל לקבל תובנות אודות מאפיינים אלו ונוכל לקבל מושג ויזואלי שיעזור להמחיש את ההשפעה שלהם על מדינת ישראל.

השאלות המשניות בהן נתרכז:

1. האם סוג היום (חול המועד, סופ"ש וכו') משפיע על כמות התאונות?
2. האם ניתן לראות קשר בין רוחב הכביש לסוגי התאונות? (שאלה זו עניינה אותנו מהפן הפסיכולוגי התנהגותי בניח האם בכבישים צרים נראה יותר תאונות חזיתיות מה שיכול להעיד על חוסר סבלנות של נהגים).
3. האם ניתן לראות קשר בין רוחב הכביש + סוג הכביש לחומרת התאונה?
4. מה כמות תאונות הדרכים שקורות באזורים שונים בארץ לפי השעות ביום?

הנתונים עליהם אנחנו מסתמכים:

קוראים לדאטה סט המרכזי "תאונות דרכים עם נפגעים – 2022". הדאטה סט הנ"ל נלקח מאתר של מאגרי מידע ממשלתיים בישראל. במקור, הנתונים הם מהאתר של הלשכה המרכזית לסטטיסטיקה. בדאטה סט הנ"ל ישנן 10405 רשומות כאשר כל רשומה הינה תיעוד של תאונה מסוימת שקרתה וכ-43 פיצ'רים כגון: סוג היום (חג / ערב חג / חול המועד / אחר), מספר כביש, סוג התאונה וכו'. אנו נבצע Feature selection על מנת למצוא את הפיצ'רים שיענו בצורה הטובה ביותר על השאלות שלנו.

עם תחילת הפרויקט ראינו כי הנתונים הנוכחיים לא יספיקו לניתוח מעמיק לצורך יצירת הויזואליזציה ולכן החלטנו לפנות לנתיבי ישראל במטרה להרחיב את הדאטה שלנו. קיבלנו דאטה סט נוסף, שיסייע לנו להרחיב את הנתונים ואת המענה לשאלות. ביצענו את המהלך להוספת הדאטה סט הנוסף לאחר ה-Preprocess שבעקבותיו ירדו המון רשומות. ב-Dataset החדש שהתווסף ישנן 11,009 רשומות. לאחר שביצענו Pre Process מעמיק על 2 Datasets ואיחדנו בניהם התקבלו 4,049 רשומות.

## 2. נתונים:

תיאור כללי של הנתונים:

משמעות העמודה	מפתחות	קרדינליות הנתונים	שם הפיצ'ר לאחר Preprocess	פיצ'ר מדאטה סט 1/ פיצ'ר מדאטה סט 2
לכל תאונה מוצמד מספר התאונה הייחודי שלה	ייחודי	3571 ערכים ייחודיים	ID	קוד תאונה / PK_TEUNA_FIKT
השנה בה אירעה התאונה	לא מפתח	[2020 – 2023]	Year	שנה / SHNAT_TEUNA
התאריך המלא בו אירעה התאונה	לא מפתח	Date Time	Date	תאריך התאונה / (יצרנו תאריך משילוב של יום, חודש ושנה)
סוג היום בו אירעה התאונה כלומר ימי עבודה סטנדרטיים, סופ"שים, חגים וערבי חג	לא מפתח	['Sunday/Thursday', 'Monday/Tuesday/Wednesday', 'Weekend/Holiday']	Day Type	סוג יום / SUG_YOM
השעה בה אירעה התאונה	לא מפתח	Day Time	Hour of Crash	שעת התאונה / SHAA
מספר הנתיבים בכביש בו אירעה התאונה	לא מפתח	['Two-way without continuous divider line', 'One-way', 'Two-way with continuous divider line', 'Unknown number of lanes', 'Rav Maslul']	Path Type	חד נתיבית / HAD_MASLUL
האם ישנן הפרעות על הכביש ואילו הפרעות	לא מפתח	['Other', 'Sand or gravel on the road', 'Dry', 'Unknown', 'Wet', 'Smeared with fuel']	Road Surface	תנאי פני מיסעה / PNE_KVISH
התנאים של הכביש בו אירעה התאונה	לא מפתח	['No defect', 'Unknown', 'Poor shoulders and bumpy road',	Road Conditions	רמת שירות / TKINUT

		'Bumpy road', 'Poor shoulders']		
רוחב הכביש בו אירעה התאונה	לא מפתח	to 10.5 meters', '10.5 to 14 7'] , 'meters', '5 to 7 meters More than 14 meters', ' ["Up to 5 meters	Road Width	רוחב הדרך / ROHAV
מה היה מזג האוויר ביום בו אירעה התאונה	לא מפתח	['Clear', 'Rainy', 'Hot', 'Foggy', 'Other']	Weather	מזג אוויר / MEZEG_AVIR
מה היו תנאי התאורה בכביש בו אירעה התאונה למשל האם היה אור יום, האם תאורת הכביש לא הייתה מותקנת וכו'.	לא מפתח	['Night with streetlights', 'Daylight', 'Night unknown', 'Night without streetlights', 'Street lights not installed', 'Limited visibility due to weather', 'Day unknown', 'twilight', 'Night with streetlights with limited visibillity']	Light Conditions	תנאי תאורה / TEURA
המחוז בארץ בו אירעה התאונה	לא מפתח	['North area', 'Central area', 'Judea and Samaria area', 'South area', 'JUdea and Samaria area', 'Jedea and Samaria area']	Area	מרחב / YEHIDA
עד כמה חמורה הייתה התאונה	לא מפתח	['Light', 'Severe', 'Fatal']	Accident Severity	חומרת תאונה / HUMRAT_TEUNA
איך פיצת התרחשה התאונה	לא מפתח	Front-to-side crash', 'Side-to-'] , 'side crash Front-to-front crash', ' , "Crash with inanimate object , 'Injury to a pedestrian' Crash with a standing ' , 'vehicle that is not parked Backward-front crash', ' , "Sliding', 'Flip Fall from a moving ' , 'vehicle	Crash Type	סוג תאונה / SUG_TEUNA

		Getting off the road or ' going up to the sidewalk', , "Other Backward-side crash', ' , "Crash with a parked vehicle ['Crash with animal'		
--	--	--	--	--

מיפוי הנתונים בהתאם לטיפולוגיה האבסטרקטית של Munzner:

<b>Dataset 1 Type</b>	Flat Table
<b>Dataset 2 Type</b>	Flat Table

<b>Attributes</b>	<b>Data type</b>
ID	Categorical
Year	Quantitative
Date	Quantitative
Day Type	Categorical
Hour of Crash	Quantitative
Path Type	Categorical
Road Surface	Categorical
Road Conditions	Categorical
Road Width	Categorical
Weather	Categorical
Light Conditions	Categorical
Area	Categorical
Accident Severity	Ordinal
Crash Type	Categorical

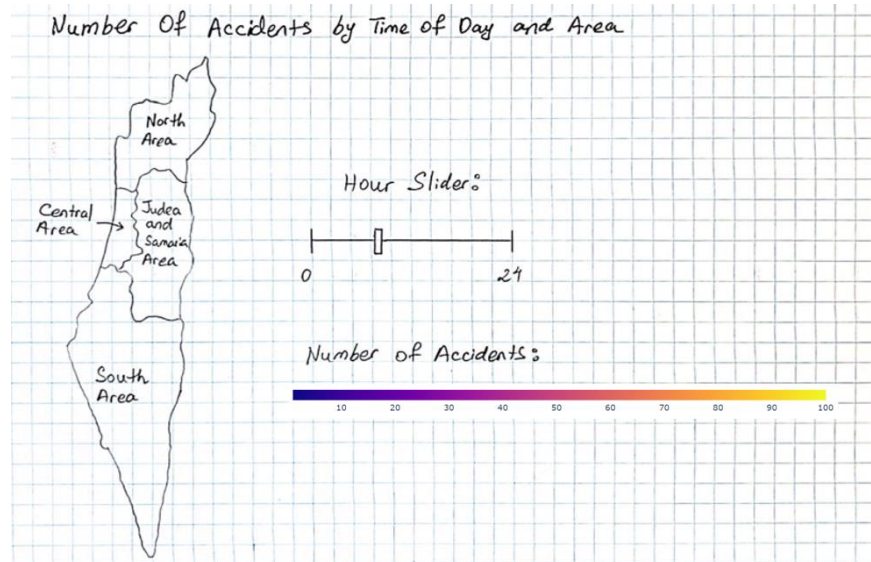
### 3. מטלות:

השאלה עליה עונים במונחי התחום	מטלות המשתמש	Target	Action
שאלה ראשית	בטיחות בדרכים	רוצים לראות התפלגות שלפיה נוכל לראות כיצד קומבינציות של מאפיינים סביבתיים משפיעים על חומרת התאונה.	השוואה
האם סוג היום (חול המועד, סופ"ש וכו') משפיע על כמות התאונות?	בטיחות בדרכים	ע"י הפעולה של מציאת הקשר נרצה לראות האם ישנה תלות בין סוג היום וכמות התאונות.	מציאת הקשר
האם ניתן לראות קשר בין רוחב הכביש לסוגי התאונות?	פסיכולוגיה	ע"י הפעולה של מציאת הקשר נרצה לראות האם ישנה תלות בין רוחב הכביש לסוגי התאונות.	מציאת הקשר
האם ניתן לראות קשר בין רוחב הכביש + סוג הכביש לחומרת התאונה?	בטיחות בדרכים	ע"י הפעולה של מציאת הקשר נרצה לראות האם ישנה תלות בין רוחב הכביש + סוג הכביש לחומרת התאונות.	מציאת הקשר
מה כמות תאונות הדרכים שקורות באזורים שונים בארץ לפי שעות ביום?	בטיחות בדרכים + דמוגרפיה	לצורך מידע ולגלות מגמה עברה נבין כיצד כמות התאונות משתנה בין אזורים שונים לאורך היום.	זיהוי

### חלק 3: עיצובים חלופיים:

חלופה 1 (מה כמות תאונות הדרכים שקורות באזורים שונים בארץ לפי השעות ביום?):

סעיף 2: סעיף א':



ב. Marks & Channels: המטרה היא להציג את כמות התאונות במחזות שונים לפי שעה ביום. ישנו slider לשעה שמאפשר למשתמש לבחור שעה ביום ולראות את כמות התאונות בכל מחוז דרך heatmap צבעוני.

Marks: **Areas** – שטחי המחזות במפת הארץ, **Lines** – הקווים המפרידים בין המחזות השונים במפת הארץ  
Channels: **Horizontal Position** – sliders של השעות, **Color** – הצבע של מפת הארץ מושפע מכמות התאונות שתקבע לכל area במפה.

ג. יתרונות של החלופה:

הערוצים הוויזואליים:

שימוש בHorizontal Position: שימוש בslider- מאפשר שליטה אינטואיטיבית ובחירה פשוטה של השעה ביום.  
שימוש בColors: ה-heatmap בצבעים מספק דרך מהירה לזיהוי אזורים עם רמות שונות של תאונות, כאשר הצבעים מאפשרים הבחנה מהירה בין האזורים השונים. בנוסף, הגיוון בצבעים עוזר לזהות תבניות או חריגות בנתוני התאונות בקלות.

אפקטיביות: Separability: השימוש בצבע ובמיקום אופקי ב-slider ברור ונפרד זה מזה, ומאפשר הבחנה קלה בין השעה ביום לבין רמות התאונות באזורים השונים.  
Popout: ה-heatmap מאפשר זיהוי מהיר של אזורים עם כמות חריגה של תאונות, כאשר צבעים שונים בצורה בולטת לעין.

Grouping: הצבעים עוזרים להבין אם יש קבוצות של אזורים עם רמות תאונות דומות. בנוסף, השימוש ב-slider עוזר לראות קשרים בין השעה ביום לבין כמות התאונות בכל אזור.

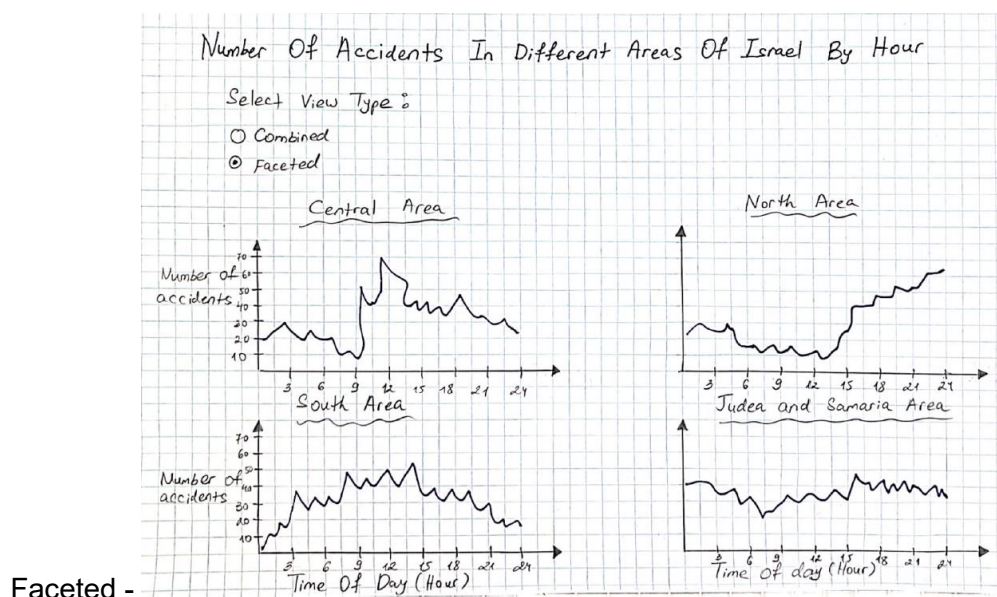
אקספרסיביות: Ordered Data - הגרף משתמש בצבעים ובHorizontal position כדי להציג נתונים כמותיים ומסודרים. לדוגמה, כמות התאונות מוצגת בצבעים שונים, כאשר כמויות גבוהות יותר של תאונות מוצגות בצבעים שונים מכמויות נמוכות יותר וגם כן slider של השעה הינו לפי סדר עולה. אלו מאפשרים להבין בצורה טובה את המגמות וההבדלים בין המחוזות והשעות.

### חסרונות של החלופה:

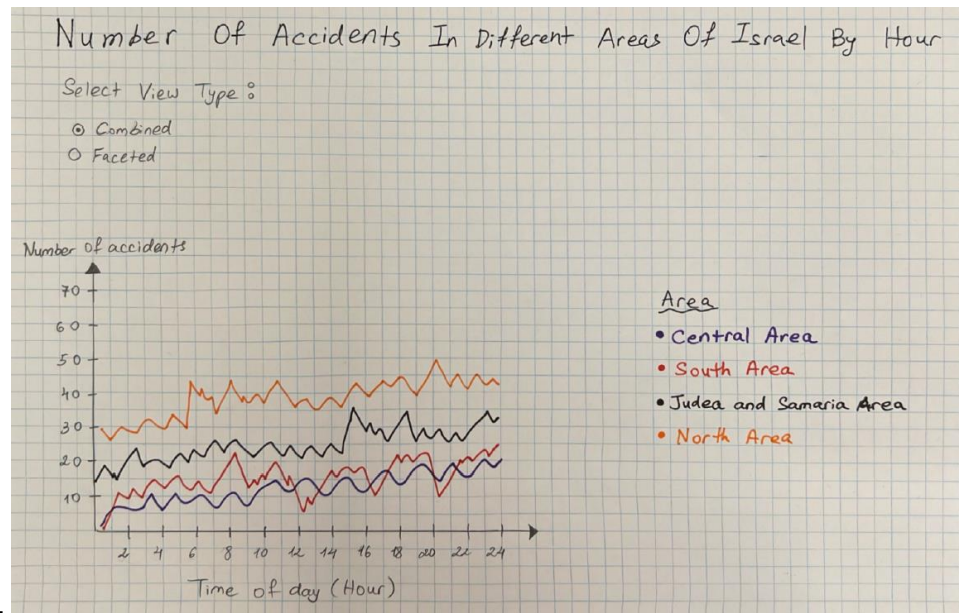
### הערצים הוויזואליים:

- שימוש במיקומים אופקיים - Horizontal Position: ה-slider מציג שעה אחת בלבד בכל פעם, ולכן קשה לראות את המגמה לאורך זמן בלי לעבור על השעות השונות ידנית.
  - שימוש בצבע - Color: ייתכן שקושי בזיהוי מדויק של כמות התאונות בהתאם לצבעים, במיוחד אם הצבעים דומים או אם יש משתמשים עם לקויות צבע.
- אפקטיביות: דיוק (Accuracy): קשה לנו להגיד באופן מדויק את ההבדל בין כמות התאונות בכל מחוז במידה והצבעים קרובים אחד לשני ובאופן כללי השימוש בסקלת הצבעים מקשה על אמירה מדויקת של כמות התאונות.
- יכולת הבחנה (Discriminability): עם השימוש בצבעים ייתכן קושי להבחין בין רמות מאוד קרובות של תאונות אם הצבעים קרובים זה לזה.
- אקספרסיביות: הערוץ של הצבע לא מותאם בצורה אידיאלית לפיצ'ר של כמות התאונות מאחר וכמות התאונות הוא מספר שלם ומפת צבעים מתאימה יותר לעבודה עם מספרים רציפים.

### החלופה בה בחרנו:







ב. **Marks & Channels**: המטרה להציג את כמות התאונות במחוזות שונים לפי השעה ביום בשתי צורות: השוואת כמות התאונות לאורך היום בין המחוזות השונים ושימוש בFaceting שמאפשר ראייה של הנתונים לפי מחוז.

#### תצוגה ללא Faceting

**Marks**: Lines – המגמות של כמות התאונות בכל מחוז  
**Channels**: Horizontal Position – השעה ביום, Vertical Position – כמות התאונות לאותה שעה, Color – לכל מחוז יש צבע ייחודי לו

#### תצוגה עם Faceting

**Marks**: Lines – המגמות של כמות התאונות בכל מחוז  
**Channels**: Horizontal Position – השעה ביום, Vertical Position – כמות התאונות לאותה שעה

#### ג. יתרונות של החלופה:

#### הערוצים הוויזואליים:

שימוש בHorizontal Position: השימוש בציר הא שיהינו ציר זמן מאפשר לנו לבצע השוואה מידית בין כמות תאונות במחוזות שונים בארץ בזמנים שונים. הדבר מתקבל באופן אופטימלי בplots ללא faceting אשר מיועד בעיקר להשוואה.

שימוש בVertical Position: השימוש בציר הy שיהינו הציר המבטא את כמות התאונות מאפשר לנו לבצע השוואה מידית בין כמות תאונות בזמנים ומחוזות שונים בארץ.

שימוש בColors רק בplots ללא Faceting: הצבעים השונים לכל מחוז מספקים דרך מהירה לזיהוי אזורים עם רמות שונות של תאונות, כאשר הצבעים מאפשרים הבחנה מהירה בין האזורים השונים.

**אפקטיביות**: נפרדות ערוצים (Separability): השימוש בצבע ובמיקום אופקי ואנכי ברור ונפרד זה מזה, ומאפשר הבחנה קלה בין השעה ביום לבין כמות התאונות באזורים השונים (בplots ללא faceting הצבע הינו

הערוץ שמייצר את ההפרדה בין האזורים ואילו Plots עם faceting מה שמייצר את ההפרדה בין האזורים זה Faceting (עצמו)

**Popout:** השימוש במיקום אנכי מאפשר זיהוי מהיר של כמות תאונות חריגה.  
**Grouping:** Plots ללא Faceting הצבעים עוזרים להבין אם יש קבוצות של אזורים עם כמות תאונות דומה. בעזרת שלושת הערוצים יחד ניתן להבחין במגמות זהות בין איזורים שונים בארץ. למשל קל לזהות בגרף ללא faceting כי יש מגמה דומה בין אזור הדרום לאזור יהודה ושומרון.

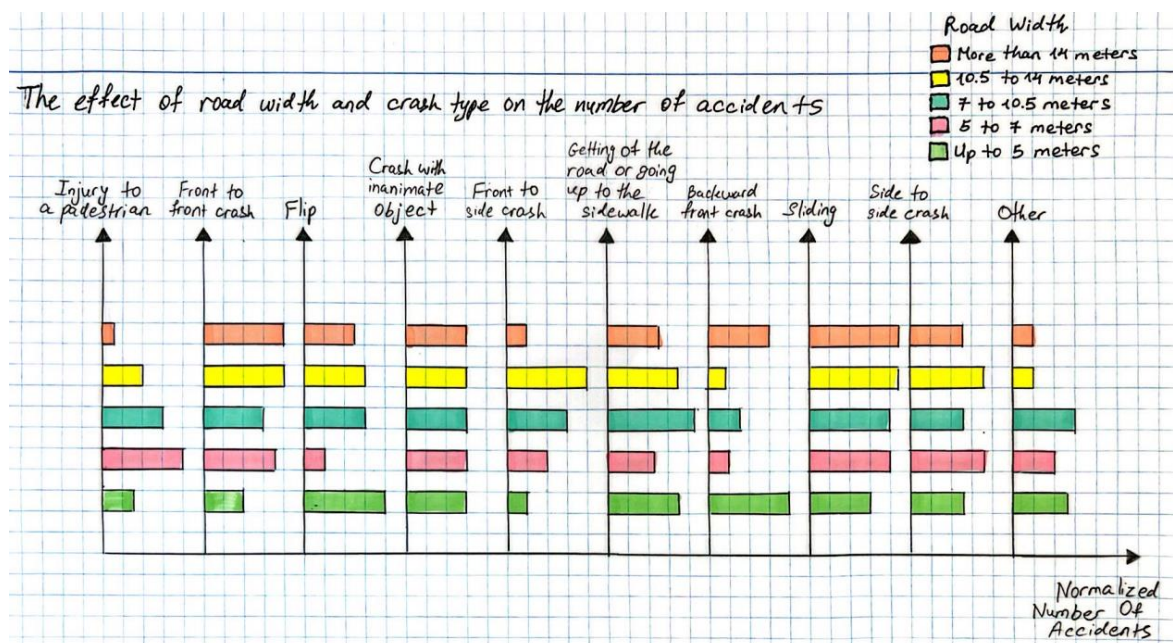
**אקספרסיביות: Ordered Data** - הגרף משתמש בHorizontal position ובVertical Position כדי להציג נתונים כמותיים ומסודרים. לדוגמה בגרף ללא Faceting כמות התאונות לאורך היום במחוזות שונים מוצגת בצבעים שונים, אלו מאפשרים להבין בצורה טובה את המגמות וההבדלים בין המחוזות והשעות. הנתונים (תאונות פר מחוז) מסודרים לפי זמן.

### חסרונות:

**ערוצים וויזואליים:** שימוש בColors בגרף ללא Faceting: אם יתקבלו מחוזות שונים בעלי מגמות זהות לחלוטין Lines שלהם יהיו אחד על השני, הדבר עלול ליצור בלבול בכך שאחד מהם יתפספס.

**הסיבה לבחירה בחלופה זו:** חלופה זו עונה בצורה מדויקת על השאלה שלנו. בנוסף, חלופה זו מאפשרת 2 צורות תצוגה, אחת שמאפשרת לנו השוואה מגמתית בין מחוזות דבר שיותר קל להבחין בו בקווי מגמה מאשר במפות חום. התצוגה השנייה משמשת להצגת נתונים בלבד פר מחוז. בחלופה זו אנו יכולים לראות את המגמה לאורך זמן בניגוד לחלופה הקודמת שמאלצת אותנו להתבייט על שעה ספציפית ביום.

### חלופה 2 (האם ניתן לראות קשר בין רוחב הכביש לסוגי התאונות?):



**ב. Marks & Channels:** המטרה להציג את כמות התאונות בכבישים ברוחב שונה לפי סוג תאונה. לכל חלק בגרף יש ציר אופקי משותף ותת גרפים עם ציר אנכי משלו. בכל תת גרף יש Bars שמייצגים רוחב כביש שונה (לפי צבע ומיקום). כל תת גרף מוגדר ע"י רוחבי כביש וסוג תאונה.

**Bars – Lines: Marks:** השונים מבטאים רוחבי כביש שונים. כל ערך של רוחב כביש מבוטא ע"י Bar במיקום וצבע קבועים.

**Channels: Horizontal Position –** ציר המציג את כמות התאונות בצורה מנורמלת לכל תת גרף. הנרמול נעשה לפי יחס התאונות ב Bar-לכמות התאונות הכללית בתת הגרף, והציר משפיע על גודל ה Bar-לפי כמות התאונות בקטגוריה,

**Vertical Position –** כאן לציר האנכי יש משמעות, הוא מגדיר את סוג התאונה בכל תת גרף ולא רק משמש למיקום ה Bars. אין חשיבות לסדר ה Bars אך יש לציין את סוג התאונה בציר האנכי.

**Color –** הצבע של כל Bar מוגדר ע"י רוחב הכביש.

### **ג. יתרונות של החלופה:**

**הערוצים הוויזואליים:** שימוש ב Colors: הצבעים השונים מחדדים מיידית את ההבדל בין ה Bars השונים ומייצרים זיהוי קל.

**אפקטיביות: Popout:** התצורה של Bar chart מייצרת השוואה נוחה ואפשרות לראות פערים גדולים בקלות.

**Grouping:** תתי הגרפים השונים מאפשרים לראות מגמות זהות בקטגוריות שונות.

**אקספרסיביות: Ordered Data:** הגרף משתמש ב Horizontal position כדי להציג נתונים כמותיים ומסודרים, בנוסף ישנו שימוש ב Vertical position וצבעים לצורך הצגת קטגוריות שונות.

### **חסרונות של החלופה:**

**הערוצים הוויזואליים:** שימוש ב Horizontal Position: ישנה בעייתיות גדולה בכך שיש ציר אופקי משותף, הדבר מייצר דחיסות בגרף ועלול לבלבל.

**שימוש ב Vertical Position:** בסידורים שונים של תתי הגרפים נקבל צורות השוואה שונות (נוחות יותר ונוחות פחות).

**אפקטיביות: דיוק (Accuracy):** קשה לנו להגיד באופן מדויק את ההבדל בין כמות התאונות בתתי גרפים שונים. כלומר אם נרצה לבצע השוואה בין אותו רוחב כביש אך עם סוג תאונה שונה יכול להיות לנו קשה.

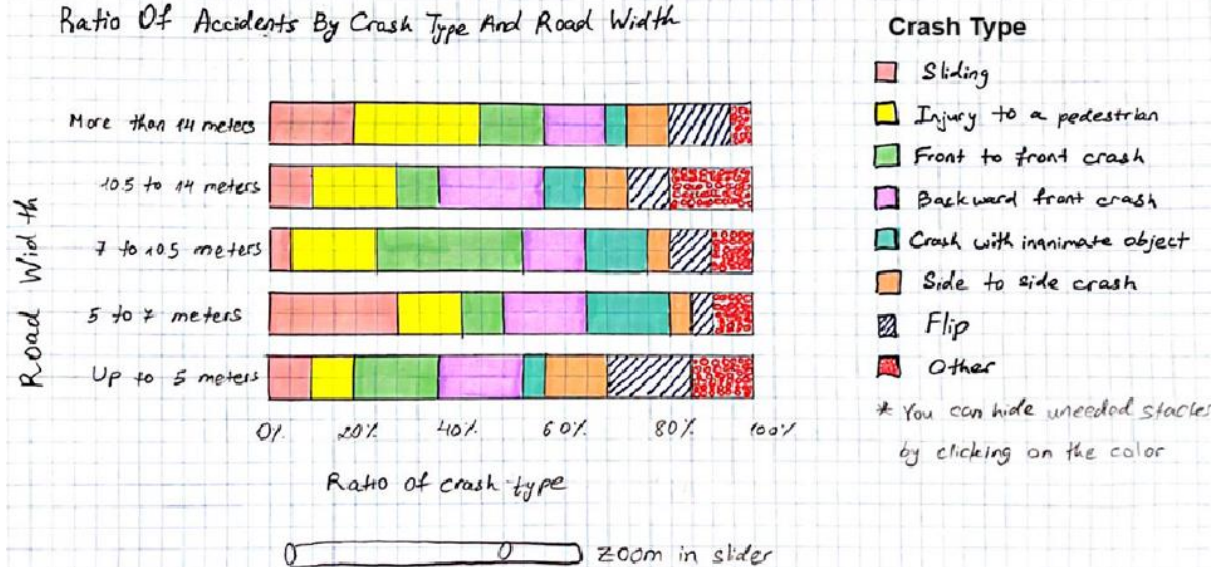
**אקספרסיביות:** הערוץ של ה Vertical position לא מותאם בצורה אידיאלית לפיצ'ר של סוג התאונה כיוון שעלול לייצר בלבול וחוסר נוחות לראות הבדלים כמו שנאמר מעלה ולכן קיים Faceting, עם זאת, בחרנו בצורה הזו כחלופה משנית מכיוון שהרבה יותר פשוט לראות מגמות זהות בעזרתה.

### **החלופה בה בחרנו:**

## Effect of Road Width On Crash Type / Accident Severity :

Search :  (select the feature you want to analyse)

Ratio Of Accidents By Crash Type And Road Width



**ב. Marks & Channels:** המטרה היא להציג את כמות התאונות בכבישים בעלי רוחב שונה לפי סוג תאונה. ניצור stacked bar כאשר כל עמודה מייצגת רוחב כביש, וכל חלק בעמודה מייצג סוג תאונה מסוים וגודל כל חלק בעמודה מייצג את כמות התאונות מאותה קטגוריה באופן יחסי. מתחת לגרף ישנו slider שמאפשר לראות רק חלק מהגרף אם אין צורך בראייה כוללת של כל סוגי התאונות – מעין zoom in.

**Marks:** Lines – ערימה אופקית של סימני קו  
**Channels:** length and color hue, spatial regions: one per glyph, Horizontal Position

**ג. יתרונות של החלופה:**

**הערוצים הוויזואליים:**

**שימוש ב-Length:** מאפשר השוואה מיידית של כמות התאונות בין סוגי תאונות שונים בכל רוחב כביש.

**שימוש ב-Color Hue:** מסייע בזיהוי והבחנה בין סוגי תאונות שונים בקלות.

**שימוש ב-Spatial Regions:** כל עמודה מייצגת רוחב כביש, מה שמאפשר הבנה ברורה של התפלגות התאונות לפי רוחב הכביש.

**שימוש ב-Horizontal Position:** נוכל לבצע zoom in על הציר האופקי בעזרת slider.

**אפקטיביות:**

**Discriminability** - קל להבחין בין סוגי התאונות השונים הודות לשימוש בצבעים שונים לכל סוג תאונה.

**Separability:** השימוש באורכי stacks והצבעים הינם ערוצים שונים כאשר הצבע מייצג את סוג התאונה ואורך הstack מייצג את כמות התאונות היחסית

**Popout:** השימוש בצבע עבור כל סוג תאונה מציג את ההבדלים בין סוגי התאונות ומאפשר זיהוי מהיר של סוגי תאונות שונים.



Grouping: השימוש בצבע מאפשר לשייך סוגי תאונות דומות לאותה הקבוצה וכן השימוש ב-spatial regions כך שכל glyph מאפשר ייצוג של קבוצות רחב כביש שונות.

### אקספרסיביות:

Categorical Data: השימוש בצבעים שונים לכל סוג תאונה עוזר לזהות בקלות את סוגי התאונות השונים שהינו ערך קטגוריאלי וכן השימוש ב-glyphs מתאר סוג דאטה שהינו קטגוריאלי – רוחב הכביש.

### חטרונות של החלופה:

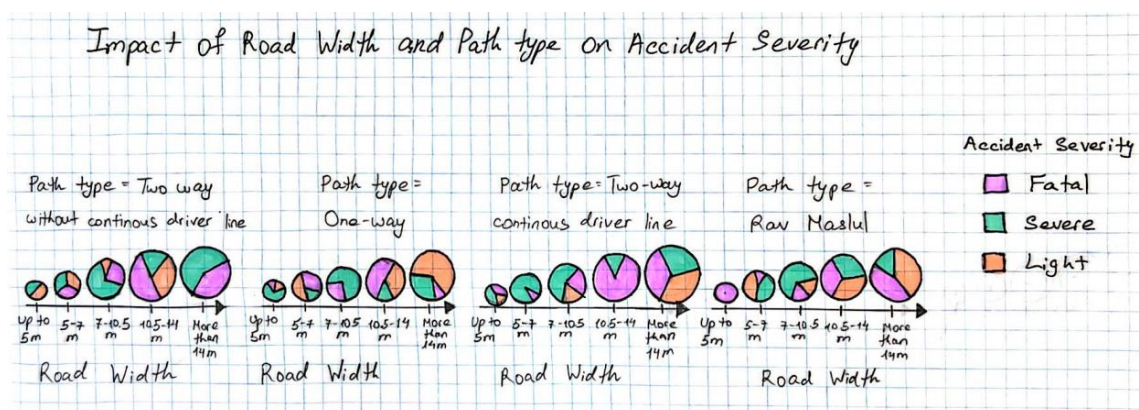
### הערוצים הוויזואליים:

Length: ייתכן וקשה להבחין בהבדלים קטנים בכמויות התאונות כאשר ה-stacks קטנים מאוד.  
Color Hue: אם יש יותר מדי קטגוריות, הצבעים יכולים להיות מבלבלים.  
One per Glyph: Spatial Regions: אם ישנם הרבה סוגי תאונות בכל רוחב כביש, ה-stacks עלולים להיות קטנים מאוד, מה שעלול לגרום לעומס ויזואלי וקושי בהבנת הנתונים.

אפקטיביות: Accuracy: השימוש בערוץ של length לא מאפשר להבחין בהבדלים קטנים בכמויות התאונות.

**הסיבה לבחירה בחלופה זו:** חלופה זו עונה בצורה מדויקת על השאלה שלנו. בנוסף, חלופה זו מאפשרת לנו לבצע השוואה בקלות בין סוגי תאונות זהות ברוחבי כביש שונים בניגוד לחלופה הראשונה שהצענו. כמו כן, השימוש ב-slide מאפשר לעשות zoom in ל-stacked bar ולראות ביתר קלות את ה-stacks השונים במצב בו יש לנו stacks מרובים וכן גם לבצע השוואה נקודתית יותר בין סוג תאונה ספציפי בין רוחבי הכביש השונים.

### חלופה 3 (האם ניתן לראות קשר בין רוחב הכביש + סוג הכביש לחומרת התאונה?):



**Marks & Channels**: המטרה להציג את כמות התאונות בכבישים ברוחב שונה לפי סוג תאונה. בחרנו לבצע מעין faceting של גרפי פאי. כל גרף מייצג path type שונה ובו כל פאי מייצג רוחב כביש שונה. בכל פאי ניתן לראות את ההתפלגות של חומרת התאונות.

Areas: Marks – גרפי הפאי השונים

**Color: Channels** – הצבע של כל חתיכה בפאי מייצגת חומרת תאונה שונה, **Size (Area)** – כל גודל פאי מציג רוחב כביש שונה, **Tilt** – כל קו בפלחי הפאי הוא בעל הטייה מסוימת.

### ג. יתרונות של החלופה:

**הערוצים הוויזואליים: Color**: השימוש בצבעים עוזר להבחין בין קטגוריות שונות של חומרת תאונות.

**אפקטיביות: Discriminability**: ערוץ הצבע מאפשר לנו להבחין בין 3 רמות שונות של תאונה בצורה מהירה וברורה. גם ערוץ size מאפשר להבחין בקלות בין רוחבי כביש שונים שמיוצגים ע"י גודל הפאי. **Separability**: השימוש בערוץ הצבע, tilt והsize הם נפרדים ומייצגים דברים שונים. הצבע – חומרת תאונה, tilt – התפלגות היחסית של כמות התאונות מחומרה מסוימת והsize שמייצג את רוחב הכביש. **Popout**: בערוץ tilt נוכל להבחין בגודל פלח שעשוי להיות חריג ביחס לפלחים אחרים בקלות. **Grouping**: ערוץ הצבע מאפשר לראות בקלות קבוצות שונות של חומרת התאונה, גם ערוץ size מאפשר לראות קבוצות של רוחבי כביש שונים בקלות.

**אקספרסיביות: size** של גדלי הפאי מתאר ordered attribute מאחר וכל רוחב כביש הינו אורדינלי, הצבעים מתארים את חומרת התאונה שהינו ערך קטגוריאלי והtilt מייצג את החלק היחסי של חומרת התאונה מתוך כלל התאונות באותה קטגוריה (היחסיות הינה ערך נומרי ולכן ordered).

### חסרונות של החלופה:

**הערוצים הוויזואליים: שימוש בTilt** – קשה להשוות באופן מדויק בין גדלי פלחים שונים, במיוחד כאשר ההבדלים קטנים. אנשים נוטים לתפוס שטחים בצורה פחות מדויקת. **Sizes** – התפיסה של ההבדלים בגדלי הפאי עשויה להיות נתונה לפרשנויות שונות אם לא משתמשים בהבדלים משמעותיים בין גדלי הפאי השונים.

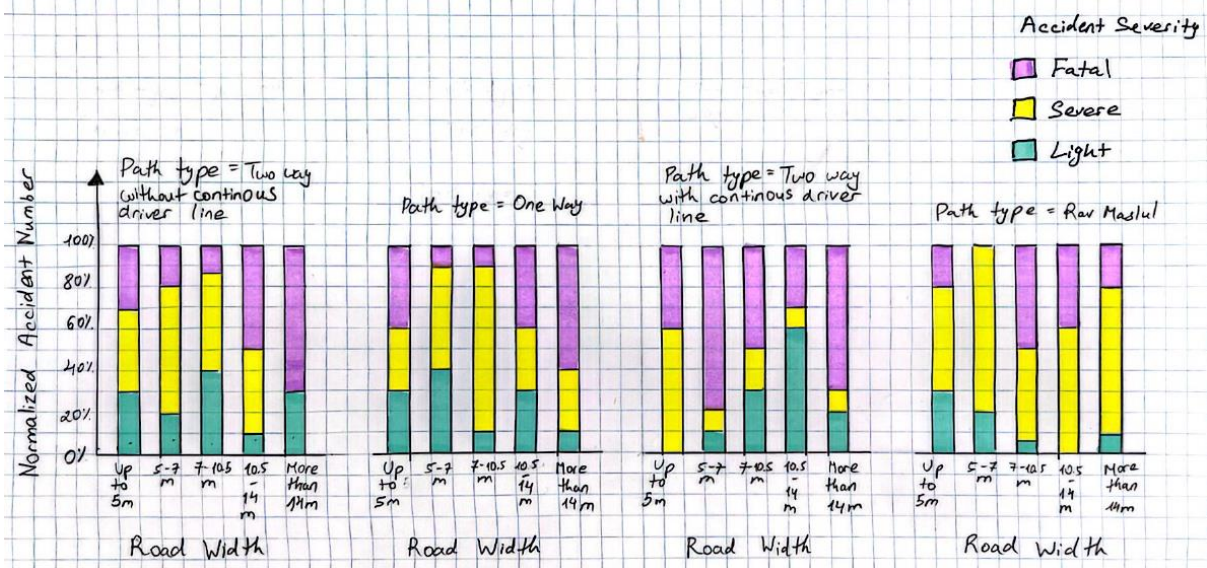
**אפקטיביות: דיוק (Accuracy)**: ערוץ tilt מקשה עלינו להגיד באופן מדויק את ההבדל בין החלק היחסי של חומרת התאונה בתוך כל פאי וגם כן ערוץ size יכול להשפיע על איך שנתפוס את החלק היחסי של חומרת התאונה כלומר בפאי קטן אנו עשויים לחשוב מאחר והפלחים באותו פאי קטנים כנראה שחומרת התאונה שם נמוכה יותר ביחס לפאי גדול יותר.

**אקספרסיביות: השימוש שלנו בציר הא לתיאור רוחב הכביש וכן בגודל פאי לצורך כך הינו מיותר, אנחנו מתארים את אותו הדבר בשני ערוצים שונים. באופן כללי קשה להסיק תובנות מהגרף הזה, הוא יחסית מורכב ולא מדויק.**

### החלופה בה בחרנו:

## Effect Of Road Width On Crash Type / Accident Severity :

Search : Accident Severity ▼ (Select the feature you want to analyze)



**ב. Marks & Channels:** המטרה להציג את כמות התאונות בכבישים ברוחב שונה לפי סוג תאונה. בחרנו לבצע faceting של גרפי Stacked Bar Charts. כל גרף מייצג path type שונה ובו כל בר מייצג רוחב כביש שונה. בכל בר ניתן לראות את ההתפלגות של חומרת התאונות.

**Marks:** Lines – ערימה אופקית של סימני (marks) קו

**Channels:** length and color hue, spatial regions: one per glyph, Horizontal Position, Vertical Position.

## ג. יתרונות של החלופה:

**הערוצים הוויזואליים:** Color: השימוש בצבעים עוזר להבחין בין קטגוריות שונות של חומרת תאונות.

**Length:** בעזרת ערוץ זה נקבל את היחס בתוך כל בר, כלומר עבור כל סוג ורוחב כביש נקבל את יחס התאונות.

**Spatial regions:** בעזרת ערוץ זה נקבל השוואה בין רוחבי כביש שונים.

**אפקטיביות:** Discriminability: ערוץ הצבע מאפשר לנו להבחין בין 3 רמות שונות של תאונה בצורה מהירה וברורה. הערוץ האופקי מאפשר להבחין בקלות בין רוחבי כביש שונים שמיוצגים ע"י הברים השונים. בעזרת הערוץ האנכי ניתן בקלות להשוות בין יחסי תאונות (מבחינת כמויות) בין חומרות תאונה זהות ברוחבי כביש שונים עקב כך שהברים ממוינים מבחינת הStacks.

**Separability:** כל ערוץ לחלוטין מציג דברים שונים. Color – חומרת תאונה, length – יחס תאונות לכל חומרת תאונה בכל רוחב כביש עבור כל סוג כביש, spatial regions – רוחבי כביש שונים.

**Popout:** בערוץ הlength נוכל להבחין בגודל Stack שעשוי להיות חריג ביחס לStacks אחרים בקלות.

**Grouping:** ערוץ הצבע מאפשר לראות בקלות קבוצות שונות של חומרת התאונה, גם ערוץ הspatial regions מאפשר לראות קבוצות של רוחבי כביש שונים בקלות.

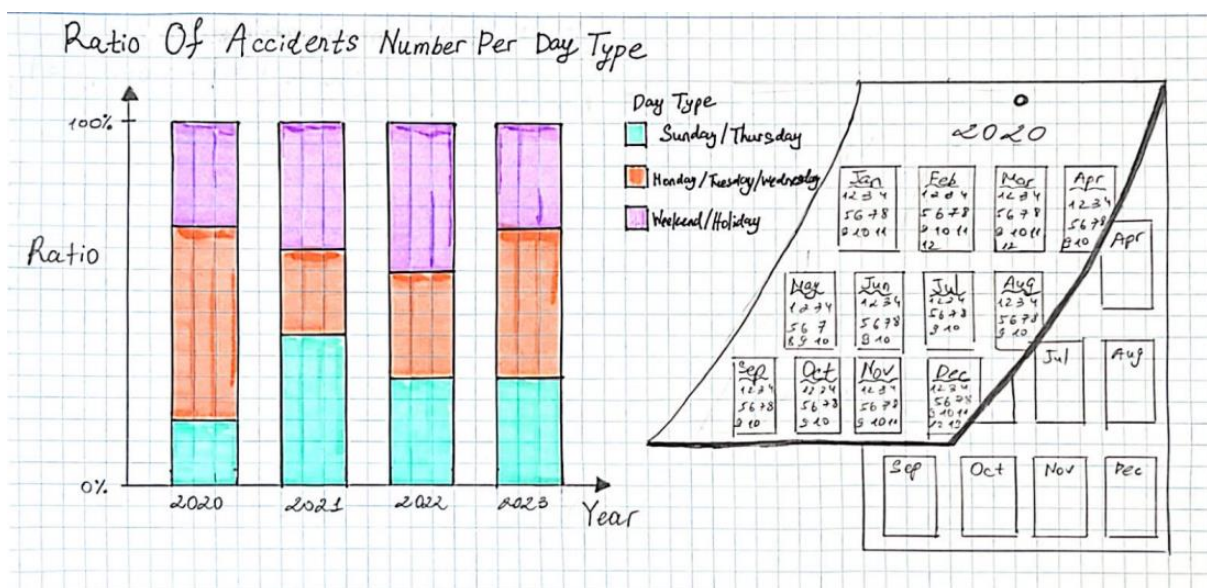
**אקספרסיביות:** ערוץ spatial regions מתאר ordered attribute מרחב וכל רוחב כביש הינו אורדינלי, הצבעים מתארים את חומרת התאונה שהינו ערך קטגוריאלי והlength מייצג את החלק היחסי של חומרת התאונה מתוך כלל התאונות באותה קטגוריה (היחסיות הינה ערך נומרי ולכן ordered).

### חסרונות של החלופה:

**אפקטיביות:** דיוק (Accuracy): ערוץ length מקשה עלינו להגיד באופן מדויק את ההבדל בין החלק היחסי של חומרת התאונה בין הברים השונים כאשר מדובר בFacets שונים אם ישנו שוני קטן ביחס.

**הסיבה לבחירה בחלופה זו:** חלופה זו עונה בצורה מדויקת על השאלה שלנו. מבחינה ויזואלית הגרף הנ"ל מאוד נוח לעין וקל להבנה ביחס לחלופה הראשונה. בנוסף, חלופה זו מאפשרת לנו לראות בצורה מדויקת ביחס לגרף הפאי את החלק היחסי מבחינת כמות תאונות של כל facet ו stack מתוך כמות התאונות הכללית של כל bar שמציין רוחב כביש בpath type מסוים.

### חלופה 4 (האם סוג היום (חול המועד, סופ"ש וכו') משפיע על כמות התאונות?):



**ב. Marks & Channels:** המטרה היא להציג את כמות התאונות לפי סוג יום בשנים שונות. ניצור stacked bar כאשר כל עמודה מייצגת שנה אחרת, וכל חלק בעמודה מייצג סוג יום מסוים וגודל כל חלק בעמודה מייצג את כמות התאונות מאותה קטגוריה באופן יחסי. לצד הגרף ישנו לוח שנה (מפולטר לפי שנה אחת כל פעם) בכדי לראות את סוג היום עבור תאריך ספציפי ובנוסף לראות את כמות התאונות באותו היום (ראייה פרטנית).

**Marks:** glyph – ברים אנכיים אשר מחולקים לתתי-ברים. Areas – שטחי לוח השנה המציג ימים באופן ספציפי.

**Channels:** Horizontal position – ציר איקס שמייצג את השנים (מחולקים לbins כך שכל שנה היא bin). Vertical position – החלק היחסי של כל תת בר מתוך סך כל התאונות מאותה קטגוריה. Length –



גובה כל תת-בר מוגדר לפי החלק היחסי של אותו סוג יום מבחינת כמות תאונות. **Color** – כל צבע מגדיר סוג יום שונה, **spatial regions: one per glyph**.

### ג. יתרונות של החלופה:

#### הערוצים הוויזואליים:

שימוש בLength: מאפשר השוואה מיידית של כמות התאונות בין סוגי יום זהים בשנים שונות.  
שימוש בColor Hue: מסייע בזיהוי והבחנה בין סוגי יום שונים בקלות.  
שימוש בSpatial Regions: כל עמודה מייצגת שנה, מה שמאפשר הבנה ברורה של התפלגות התאונות לפי שנים.  
שימוש בHorizontal Position: נותן הפרדה ברורה בין השנים השונות.

#### אפקטיביות:

Discriminability - קל להבחין בין סוגי הימים השונים הודות לשימוש בצבעים שונים לכל סוג תאונה ובנוסף קל להבחין בין השנים השונות לפי הברים השונים.  
Separability: יש הפרדה מלאה בין הערוצים, כל ערוץ מציג מידע שאינו מוצג בערוץ אחר.  
Grouping: השימוש בצבע מאפשר לשייך סוגי תאונות דומות לאותה הקבוצה וכן השימוש בspatial regions כך שכל glyph מאפשר ייצוג של קבוצות לפי שנים שונות.

### חסרונות של החלופה:

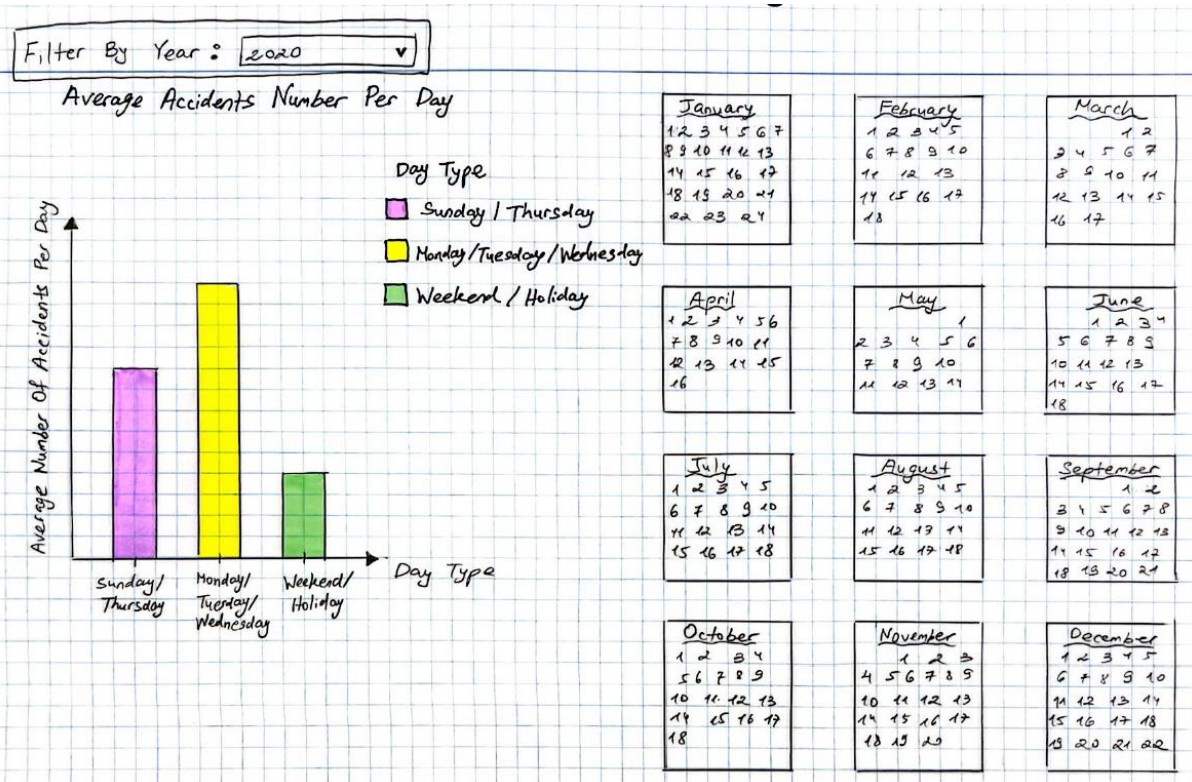
#### הערוצים הוויזואליים:

Length: ייתכן וקשה להבחין בהבדלים קטנים בכמויות התאונות כאשר ה-stacks קטנים מאוד.  
Color Hue: אם יש יותר מדי קטגוריות, הצבעים יכולים להיות מבלבלים.  
One per Glyph: Spatial Regions: אם ישנם הרבה שנים בדאטה, עלולים להתקבל הרבה ברים. בנוסף אם יהיו הרבה סוגי יום נקבל הרבה stacks ובכך נאבד מההבנה הפשוטה.  
Area: נוצרת בעיית memory הלוח שנה שמוצג לצד הגרף מראה שנה אחת בכל פעם, אם נרצה להסתכל על הגרף ועל הלוח שנה במקביל לא נוכל להצליב את כל המידע בבת אחת (בגרף מוצגות כל השנים ואילו בלוח שנה אחת ועל כן הבעיה).  
Vertical Position: שימוש ב- לא קל לראות מה החלק היחסי של כל קטגוריה.

אפקטיביות: Accuracy: השימוש בערוץ של length לא מאפשר להבחין בהבדלים קטנים בכמויות התאונות.

אקספרסיביות: המידע לא מוצג בצורה שווה מבחינת הערוצים – בציר האופקי אנחנו משתמשים בשנים כמשתנה קטגוריאלי הפורס את כל המידע ואילו בArea (לוח השנה) אנחנו מציגים את המידע פר שנה ספציפית וכך מתקבל חוסר איזון בהצגת המידע.

### החלופה בה בחרנו:



**ב. Marks & Channels:** המטרה היא להציג את כמות התאונות לפי סוג יום. ניצור bar chart כאשר כל עמודה מייצגת סוג יום אחר. הגובה של כל בר מציג את כמות התאונות הממוצעת לאותו סוג יום. לצד הגרף ישנו לוח שנה בכדי לראות את סוג היום עבור תאריך ספציפי ובנוסף לראות את כמות התאונות באותו היום (ראייה פרטנית). הגרף מפולטר כולו (כולל לוח השנה) לפי שנה ספציפית).

**Marks:** lines – ברים אנכיים. Areas – לוח השנה המציג ימים באופן ספציפי, לוח השנה מורכב מ-12 צורות.

**Channels:** Horizontal position – ציר איקס שמייצג את סוגי הימים. Vertical position – כמות התאונות הממוצעת של כל סוג יום באותה שנה. Length – גובה כל בר מוגדר לפי כמות התאונות הממוצעת של אותו סוג יום. Color – כל צבע מגדיר סוג יום שונה.

**ג. יתרונות של החלופה:**

**הערוצים הוויזואליים:**

**שימוש ב-Length:** מאפשר השוואה מיידית של כמות התאונות בין סוגי יום שונים.

**שימוש ב-Color Hue:** מסייע בזיהוי והבחנה בין סוגי יום שונים בקלות.

**שימוש ב-Horizontal Position:** נותן הפרדה ברורה בין סוגי היום השונים.

**שימוש ב-Vertical Position:** נותן הבחנה מהירה וקלה של כמות התאונות לפי אותו סוג יום. אין צורך בלהתחיל לבדוק חלק יחסי.

**אפקטיביות:**

**Discriminability** – קל להבחין בין סוגי הימים השונים הודות לשימוש בצבעים שונים לכל סוג תאונה.

**Separability:** יש הפרדה כמעט מלאה בין הערוצים, כל ערוץ מציג מידע שאינו מוצג בערוץ אחר מלבד הערוץ האופקי וערוץ הצבע – שניהם מבטאים את סוג היום אך השילוב של שניהם דווקא עוזר כיוון שמתקבל מיון (ציר X) בנוסף להפרדה מבחינת הקטגוריות (צבעים).

**Grouping:** השימוש בציר האופקי מאפשר לראות מגמות דומות בסוגי יום שונים. השימוש ב-Area (לוח שנה) מאפשר זאת בצורה פרטנית.

**Accuracy:** Vertical Positionn מאפשר לנו לראות בצורה מדויקת (יחסית) את כמות התאונות הממוצעת פר סוג יום.

**Popout:** בשימוש בערוץ הצבע והציר האנכי נוכל להבחין בסוגי יום חריגים ביחס לאחרים.

**אקספרסיביות:** אנו משתמשים בערוץ הצבע לערכי סוג יום (קטגוריה), ערוץ הגובה עבור כמות תאונות ממוצעת (נומרי).

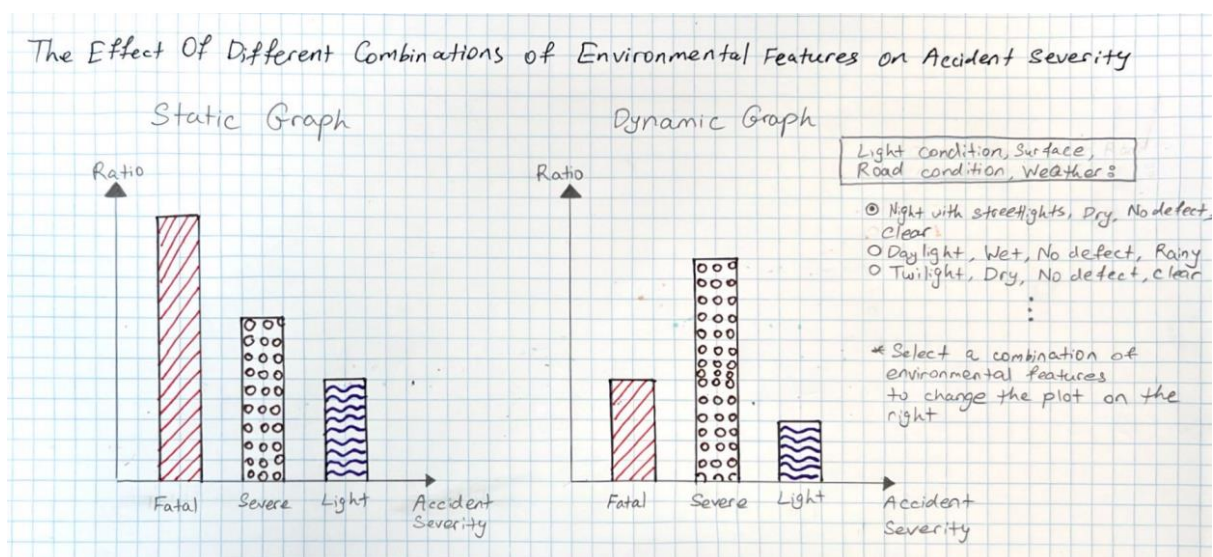
### חסרונות של החלופה:

#### הערוצים הוויזואליים:

**Area:** בכדי לראות את כל לוח השנה צריך לגלול למטה כך שהגרף כבר לא מולנו, הבעיה הזו היא בעיה טכנית עבורנו ואם היינו יודעים לתכנת בשפת עיצוב היינו פותרים אותה.

**הסיבה לבחירה בחלופה זו:** חלופה זו עונה בצורה מדויקת על השאלה שלנו. בנוסף, השימוש ב-bar chart הוא פשוט ויותר נכון ביחס לחלופה הראשונה בכך שהפיצ'ר של השנים כעת כבר אינו רלוונטי (עדיין ניתן לפלטר את כל הגרף לפי שנים). בחלופה הראשונה הגרף מציג מידע על פני מספר שנים ואילו לוח השנה מציג מידע על פני שנה אחת בלבד. בחלופה שבה אנחנו בחרנו יש תיאום בין bar chart ללוח השנה וכך האקספרסיביות גוברת.

### חלופה 5 (כיצד מאפיינים סביבתיים שונים השפיעו על תאונות הדרכים בישראל):



ב. **Marks & Channels**: המטרה היא להשוות בין תנאים סביבתיים (קומבינציות) שונים כאשר משמאל מוצג גרף קבוע (קומבינציה אידאלית של תנאים) ומימין ישנו גרף דינאמי שמשמש לשם ההשוואה אל מול הסטטי. נרצה לראות כיצד הם משפיעים על כמות התאונות ביחס לחומרת התאונה.

**Marks**: lines – ברים אנכיים.

**Channels**: **Horizontal position** – ציר איקס שמייצג את חומרת התאונה. **Vertical position** – החלק היחסי של כל בר מתוך סך כל התאונות מאותה קומבינציה. **Length** – גובה כל בר מוגדר לפי החלק היחסי של אותו סוג תאונה מבחינת כמות תאונות. **Color** – כל צבע מגדיר חומרת תאונה שונה.

### ג. יתרונות של החלופה:

#### הערוצים הוויזואליים:

שימוש בColors: מסייע בזיהוי והבחנה בין חומרות תאונה שונות בקלות.

שימוש בHorizontal Position: נותן הפרדה ברורה בין חומרות התאונה השונות ומיון לפי קשה-קל.

#### אפקטיביות:

Accuracy – נוכל לראות בצורה מדויקת מה היחס תאונות עבור חומרת תאונה מסוימת בעזרת הציר האנכי.

Discriminability – קל להבחין בין חומרות תאונה שונות הודות לשימוש בערוץ הצבע.

Separability: יש הפרדה כמעט מלאה בין הערוצים, כל ערוץ מציג מידע שאינו מוצג בערוץ אחר מלבד

הערוץ האופקי וערוץ הצבע אשר מבטאים את חומרת התאונה.

Grouping: השימוש בצבע מאפשר לשייך חומרות תאונות דומות לאותה הקבוצה

Popout: ע"י שימוש בציר האנכי ובצבע ניתן להבחין בחריגים ביתר קלות.

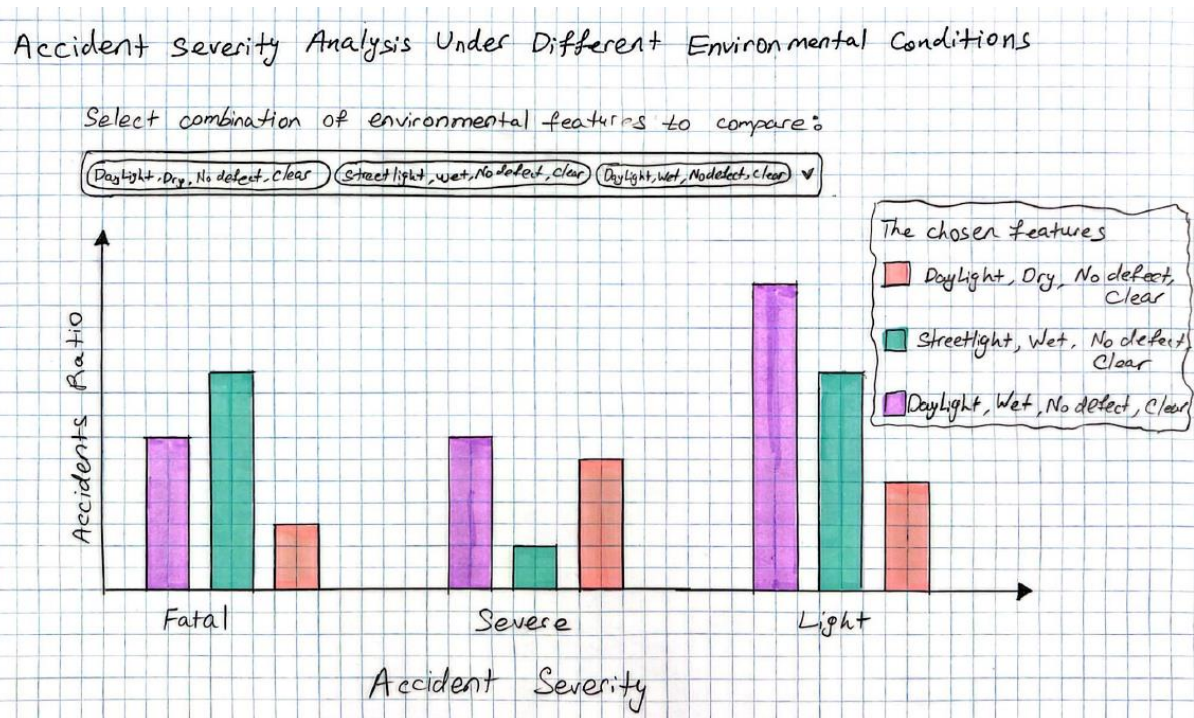
אקספרסיביות: ערוץ הצבע מתאים לחומרת התאונה (קטגוריה), ערוץ הגובה מתאים ליחס התאונות (נומרי).

#### חסרונות של החלופה:

הערוצים הוויזואליים: שימוש בLength: לא מאפשר השוואה מיידית של יחס התאונות בין חומרות תאונה

זהות בקומבינציות שונות.

#### החלופה בה בחרנו:



**ב. Marks & Channels:** המטרה היא להשוות בין תנאים סביבתיים (קומבינציות) שונים וכיצד הם משפיעים על כמות התאונות ביחס לחומרת התאונה.

**Marks:** lines – ברים אנכיים.

**Channels:** **Horizontal position** – ציר איקס שמייצג את חומרת התאונה. **Vertical position** – החלק היחסי של כל בר מתוך סך כל התאונות מאותה קומבינציה. **Length** – גובה כל בר מוגדר לפי החלק היחסי של אותו סוג תאונה מבחינת כמות תאונות. **Color** – כל צבע מגדיר חומרת תאונה שונה.

**ג. יתרונות של החלופה:**

**הערוצים הוויזואליים:**

**Colors:** שימוש בזהוי והבחנה בין חומרות תאונה שונות בקלות.

**Horizontal Position:** שימוש בפרדה ברורה בין חומרות התאונה השונות ומיון לפי קשה-קל.

**אפקטיביות:**

**Accuracy** – נוכל לראות בצורה מדויקת מה היחס תאונות עבור חומרת תאונה מסוימת בעזרת הציר האנכי.

**Discriminability** – קל להבחין בין חומרות תאונה שונות הודות לשימוש בערוץ הצבע.

**Separability:** יש הפרדה מלאה בין הערוצים, בשונה מהחלופה הקודמת כאן הציר האופקי אינו משמש לאותו תפקיד כמו ערוץ הצבע וזאת מכיוון שהצבע מפריד בין קומבינציות והציר האופקי מבטא את חומרת התאונה.

**Grouping:** השימוש בערוץ האופקי מאפשר לנו לבצע חלוקה בצורה פשוטה לקבוצות לפי חומרת תאונה.

בנוסף ערוץ הצבע מקבץ לקבוצות באופן ישיר, הקבוצות נקבעות לפי הקומבינציות השונות.

**Popout:** ע"י שימוש בציר האנכי ובצבע ניתן להבחין בחריגים ביתר קלות.

**אקספרסיביות:** ערוץ הצבע מתאים לקומבינציות השונות (קטגוריה), ערוץ הגובה מתאים ליחס התאונות (נומרי).

### **חסרונות של החלופה:**

כאשר יש הרבה קומבינציות בגרף עלולה להיווצר בעיה בלהשוות בין הברים השונים עקב עומס מול העיניים.

**הסיבה לבחירה בחלופה זו:** חלופה זו עונה בצורה מדויקת על השאלה שלנו. בנוסף, אנחנו יכולים לבצע השוואה בין כל הקומבינציות אל מול הקומבינציה האידיאלית בבת אחת בניגוד לחלופה הראשונה אותה הצגנו. מתוך כך אנחנו יכולים לקבל תמונה כללית רחבה יותר על התנאים הסביבתיים האידיאליים / הלא אידיאליים.

## חלק 4: הסבר על העיצוב שנבחר ועל יישומו:

### סעיף א' – Pre Process:

כבר עם תחילת העבודה הבנו כי datasetn שבחרנו מאתר Gov מאתגר מאוד, כל העמודות והנתונים היו בעברית. בשלב הראשוני בכדי להקל על העבודה העתידית תרגמנו את הכל לאנגלית. מה שעשינו היה לעבור על כל עמודה ולייצר mapping עבור הערכים שלה בעברית לערכים באנגלית. לאחר התרגום ביצענו הורדת רשומות עם ערכים חסרים/לא רלוונטיים, גילינו כי כמות הרשומות ירדה משמעותית ולכן פנינו לנתיבי ישראל במטרה להרחיב את הנתונים. נענינו בחיוב וקיבלנו dataset נוסף אשר עוקב אחר אותה מטרה (כמות תאונות בישראל) אך לא כל העמודות היו בעלות משמעות זהה, בנוסף dataset זה גם כן הגיע בעברית הן מבחינת שמות עמודות והן מבחינת רשומות. ביצענו את תהליך התרגום גם כאן ולאחר מכן ניסינו להבין כיצד לאחד בין datasets. ביצענו feature selection על כל אחד מהdatasets במטרה למצוא עמודות בעלות משמעות זהה ולבסוף קיבלנו dataset מאוחד. בדאטה המאוחד הייתה עמודה בשם Day type אשר מסבירה את סוג היום (Weekday, Holiday, Holiday Eve, Other), בעמודה זו לא מצאנו כל כך הרבה משמעות ועל כן יצרנו עמודה חדשה בשם Day type אשר מסבירה את סוג היום אחרת (Sunday/Thursday, Weekend/Holiday, Monday/Tuesday/Wednesday) עומס בכביש בצורה שונה. יצירת העמודה התבססה על עמודה אחרת בשם Day in the week אשר הינה עמודה מספרית שמבטאת את מספר היום (1 = ראשון וכן הלאה..) והשוואה עם העמודה Day type המקורית לזיהוי אילו ימים הינם ימי חג (ערבי חג נחשבים כימי חג גם כן).

### סעיף ב':

ויזואליזציה שונה שענתה על השאלה "כיצד מאפיינים סביבתיים שונים משפיעים על תאונות הדרכים בישראל?" – תחילה רצינו לענות על השאלה הראשית בעזרת Scatter plot ולכן חיפשנו משתנה רציף לציר ה-X, הדבר הוביל לכך שבציר ה-X היה פיצ'ר של KM (מיקום התאונה לפי קילומטר בכביש), ציר ה-Y היה כמות תאונות, צבע הנקודה היה סוג התאונה ובנוסף היו פילטרים למס' כביש והמאפיינים הסביבתיים (כמו שמופיע כעת בויזואליזציה הסופית). גילינו כי הגרף אינו אינפורמטיבי עקב כמות נקודות קטנה (לאחר פילטור לפי מס' כביש) ולאחר מכן הגענו לווריאציה שלדעתנו עונה יותר טוב על השאלה (הויזואליזציה הסופית).

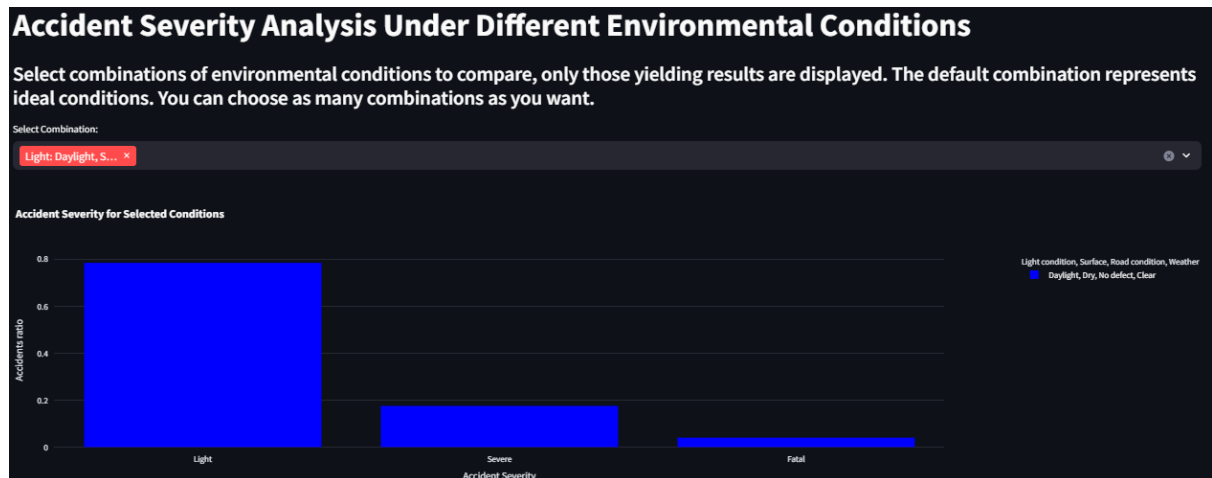
ויזואליזציה שונה שענתה על השאלה "כיצד סוג היום משפיע על כמות התאונות?" – Datasetn הגיע עם פיצ'ר בשם Day type אשר ביטא את סוג היום לפי הערכים Weekday, Holiday, Holiday Eve, Other, תוך כדי עבודה הבנו שעדיף לנו לחלק את סוג היום ל-3 קטגוריות לפי עומסי התנועה בכביש ובכך הגענו לקטלוג הבא: Sunday/Thursday, Monday/Tuesday/Wednesday, Weekend/Holiday אשר עם ערכים אלו אנו מרגישים יותר טוב בקשר לשאלה.

### סעיף ג':

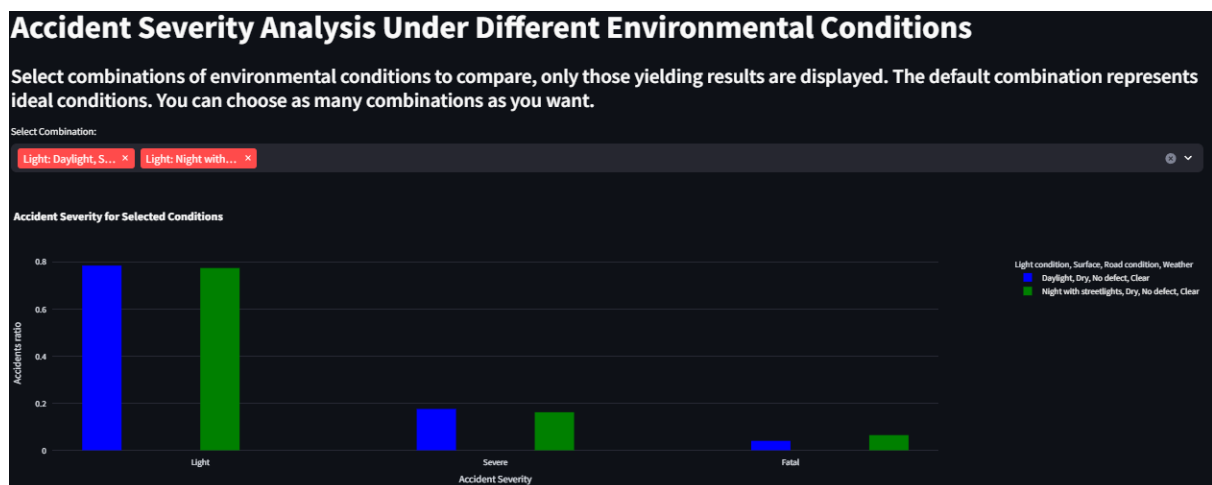
זוהי הויזואליזציה של השאלה הראשית שלנו. למעלה יש לנו מעין סרגל של פילטור שמאפשר לנו לבחור קומבינציה מסוימת של תנאים סביבתיים שמשפיעים על הכמות היחסית של התאונות בהתאם לחומרת



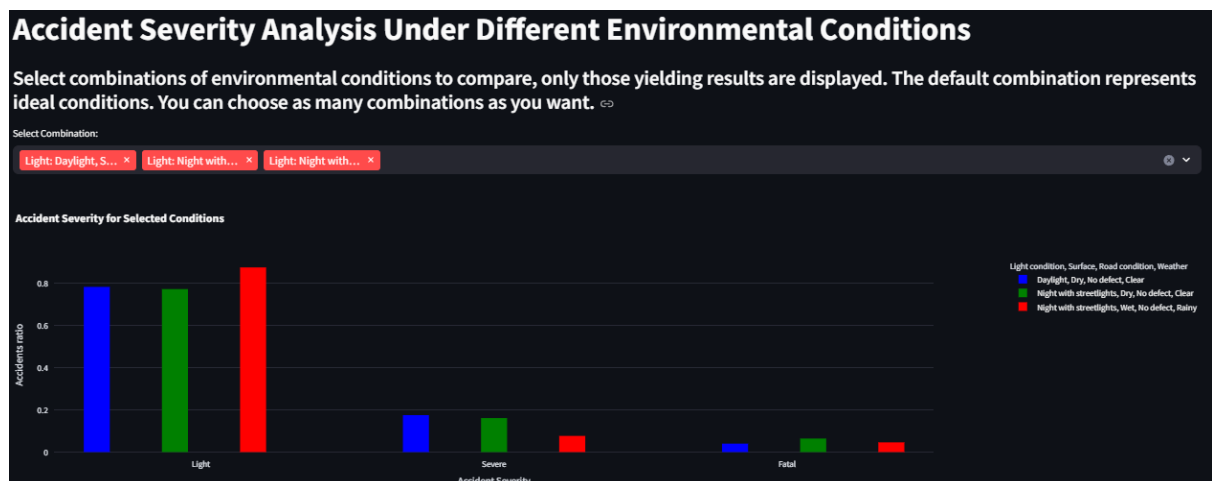
התאונה. כאשר אנחנו נכנסים לוויזואליזציה אנו מקבלים את הגרף הבא שהוא מחזיק את התנאים שהיו נראים בעינו כאידיאליים – Daylight, Dry, No defect, Clear:



ולאחר מכן אנחנו מבצעים את ההשוואה של התנאים הסביבתיים אל מול הקומבינציה האידיאלית הנ"ל, נבחר בפלטור נוסף (קומבינציה של תנאים נוספת):

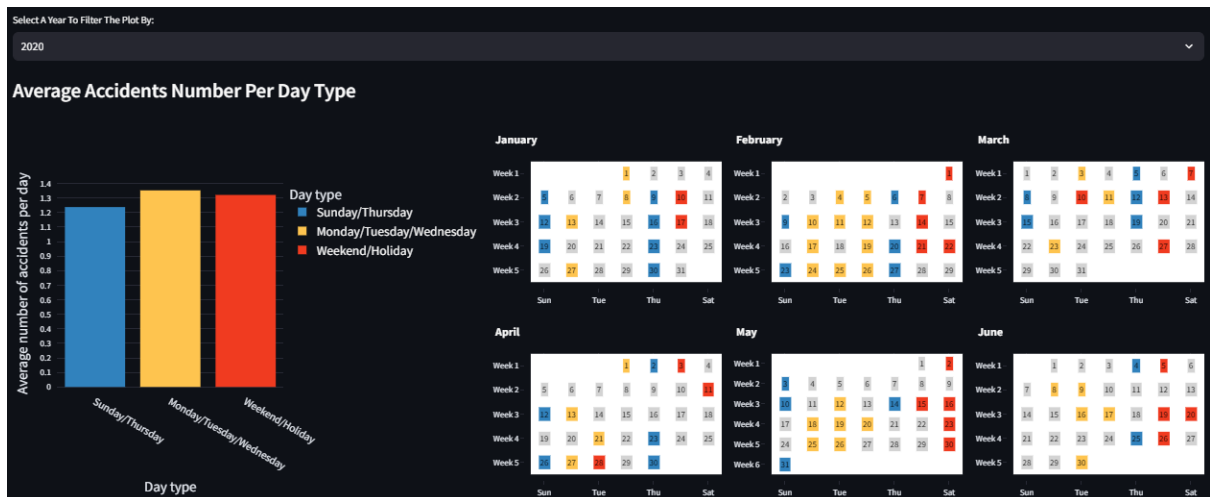


קומבינציית תנאים נוספת:

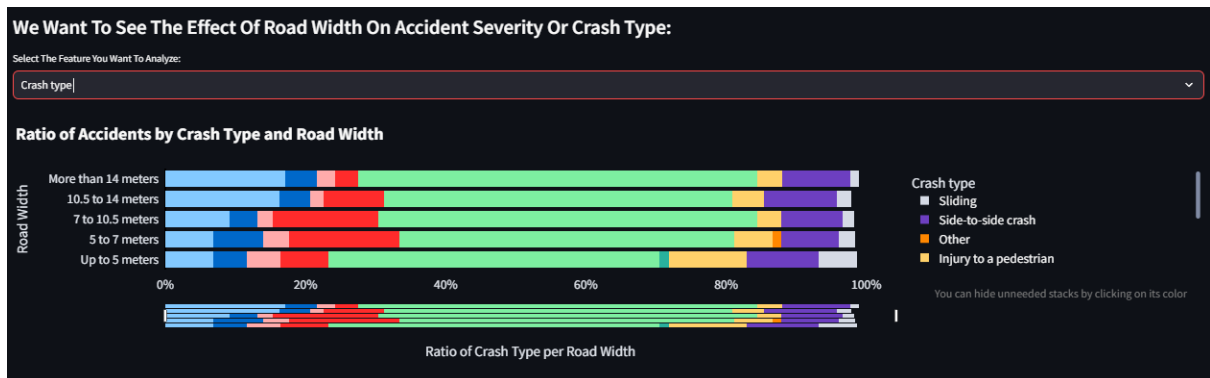




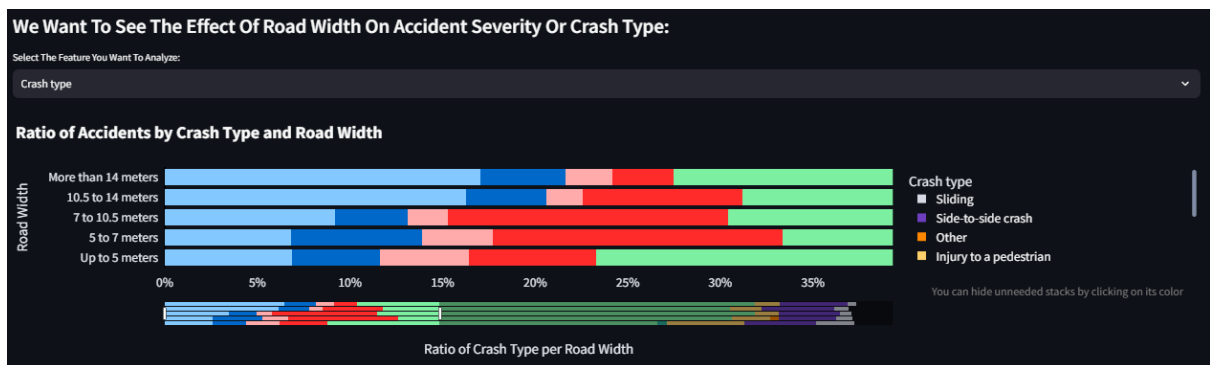
ויזואליזציה עבור השאלה - האם סוג היום (חול המועד, סופ"ש וכו') משפיע על כמות התאונות? למעשה יש לנו למעלה סרגל של פלטור שמאפשר לפלטר לפי שנים. משמאל יש לנו bar plot כך שכל צבע בו מופיע גם בלוח השנה מימין. למשל הצבע הכחול שמתאר את ימי ראשון וחמישי גם כן ייצבע בלוח שנה בהתאמה. נוכל להעביר את העכבר על לוח השנה ונוכל לראות עבור יום מסוים באופן ספציפי מה הייתה כמות התאונות באותו היום ומה היה סוג היום. אם נגלול מטה נוכל לראות את שארי חודשי השנה:



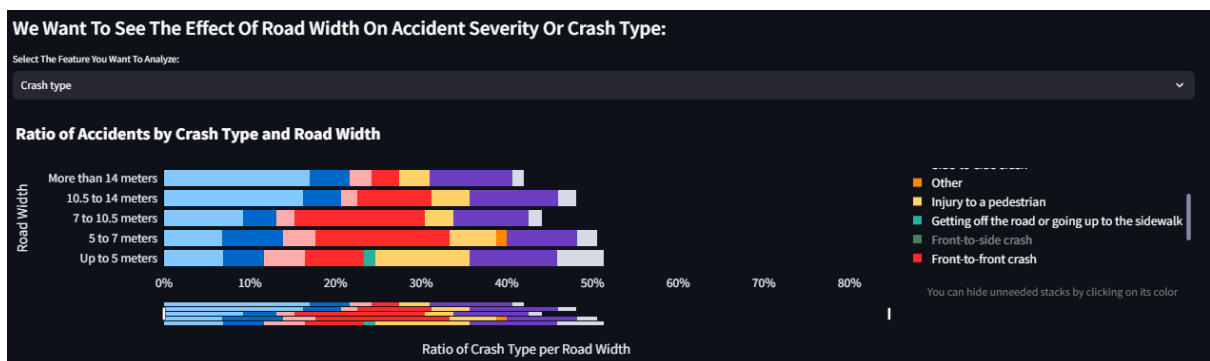
ויזואליזציה עבור השאלה - האם ניתן לראות קשר בין רוחב הכביש לסוגי התאונות? יש לפנינו stacked bar  
 כאשר כל bar מתאר רוחב כביש שונה, מימין יש לנו legend עם צבעים שמתאר סוג שונה של תאונה שקרתה:



מתחת לbar stacked השתמשנו במעין slider שיאפשר לנו לעשות zoom in כדי להתמקד באופן ספציפי בסוג תאונה מסוימת ולראות את ההבדלים עבורה בין רוחבי הכביש השונים:

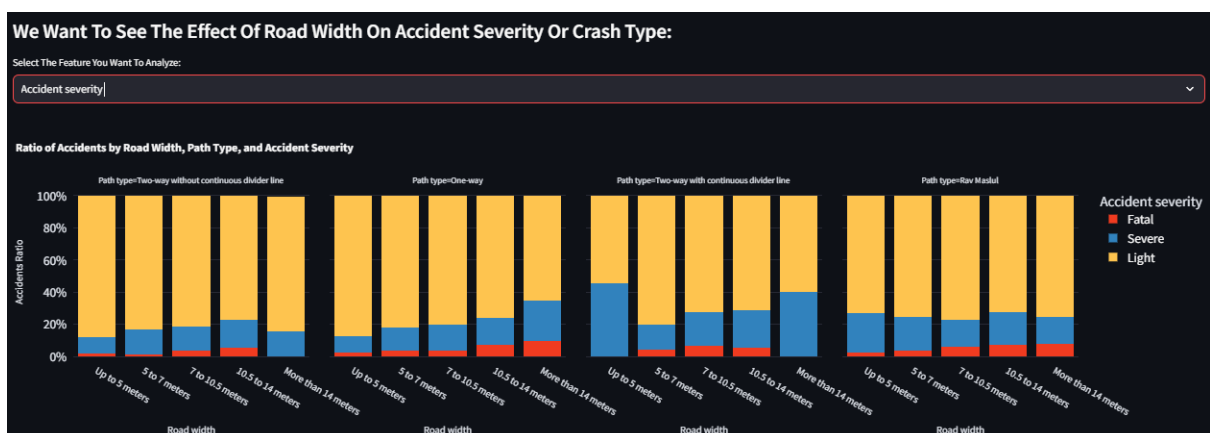


ישנה גם אפשרות לבצע לחיצה על סוג תאונה מסוים מה שיביא להורדת סוג התאונה הנ"ל מהבארה stacked כלהלן (ביצענו הורדה של Front to side crash):

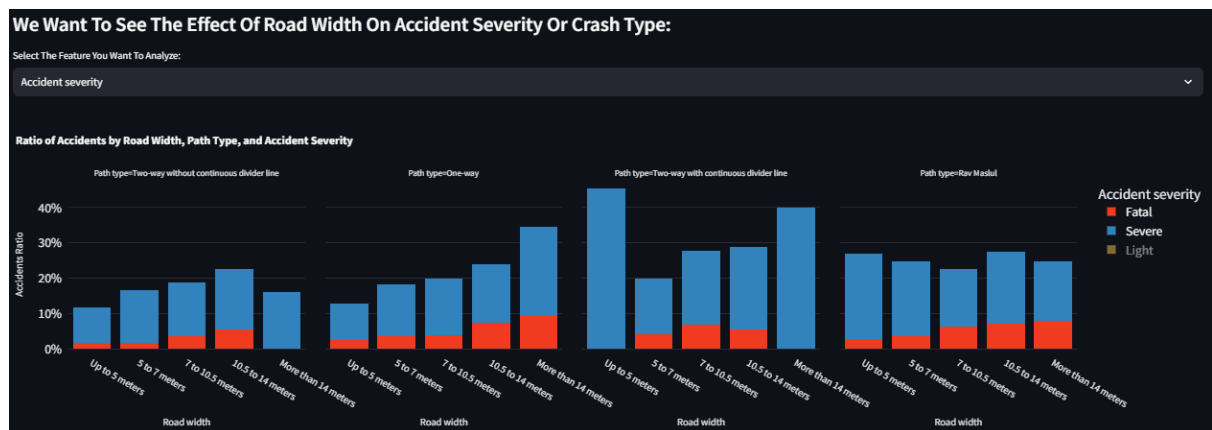


למעלה השתמשנו בסרגל פילטור שיעביר אותנו לשאלה הבאה מאחר ושתי השאלות מתייחסות לרוחב הכביש, נעבור כעת לשאלה זו.

ויזואליזציה עבור השאלה - האם ניתן לראות קשר בין רוחב הכביש + סוג הכביש לחומרת התאונה? ביצענו faceting כאשר כל תת גרף מתאר סוג מסוים של דרך. bars מתארים רוחבי כביש שונים עבור סוג דרך מסוימת. בתוך כל bar יש stacks שמתארים את הכמות היחסית של התאונות בהתאם לרמת החומרה שלהן כאשר בחרנו לשים את החומר הקטלנית ביותר כstack הראשון ולצבוע אותה באדום ואת החומרה הקלה ביותר הצבנו למעלה וצבענו אותה בצהוב שהינו צבע יותר ניטרלי:

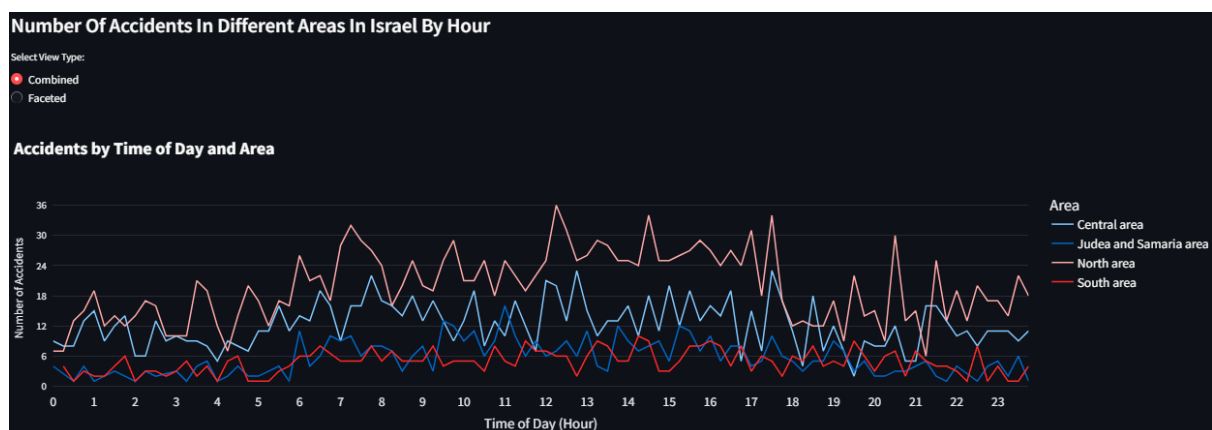


גם כאן יש לנו אפשרות ללחוץ על חומרת תאונה מסוימת והיא תעלם למשל אם נרצה להדגיש את החשיבות של התאונה הקטלנית ותאונות severe:

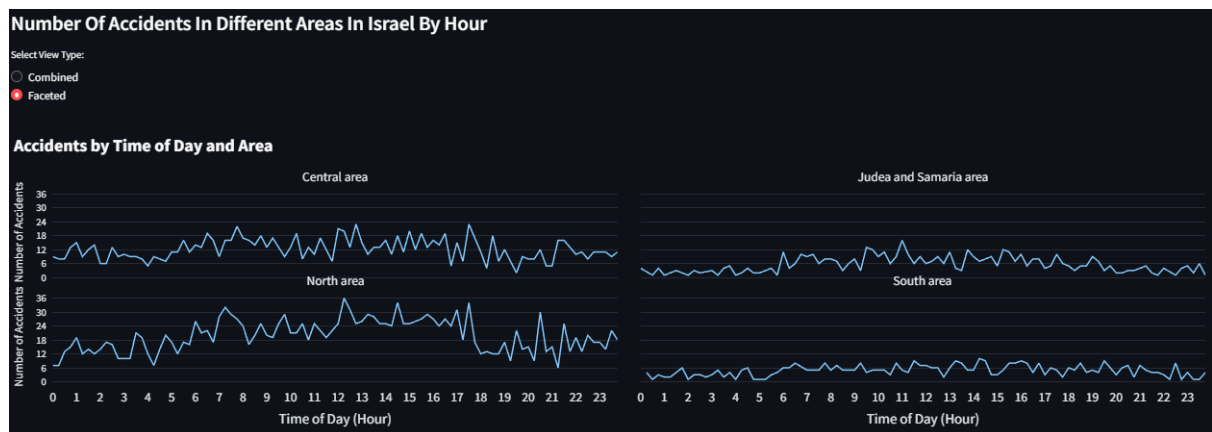


ויזואליזציה שעונה על השאלה - מה כמות תאונות הדרכים שקורות באזורים שונים בארץ לפי שעות ביום?

למעשה יצרנו עבור שאלה זו 2 צורות תצוגה – משמאל ניתן לבחור בצורת התצוגה הרצויה. תצוגה אחת שהינו משולבת, יש איחוד של כל המגמות עבור כלל מחוזי הארץ על מנת שנוכל לבצע השוואה בצורה שתהיה נוחה למשתמש. מימין יש לנו legend של צבע שמתאר מחוז מסוים. יש אפשרות לעבור על מגמה מסוימת עם העכבר ונקבל באופן ספציפי את האיזור, השעה בה אירעה התאונה ומספר התאונות באותה השעה. גם כאן יש לנו אפשרות ללחוץ על צבע בlegend והמגמה תעלה:



צורת תצוגה שנייה עבור אותה שאלה היא באמצעות faceting, גם כאן נוכל לעבור עם העכבר על המגמה ולקבל באופן נקודתי את האזור, כמות התאונות והשעה בה אירעה התאונה:



## סעיף ד':

בכתיבת הקוד השתמשנו בספריות הבאות: `base64`, `calendar`, `pandas`, `plotly`, `streamlit`. בכתיבת הקוד ל-Dashboard שלנו לא השתמשנו בשום דוגמא/קוד קיים, הכל פרי יצירתנו. הרעיון התחיל בכך שבחרנו ב-Datasets שברעיון עניין אותנו וישירות עלתה לנו שאלה ממנו. לאחר מכן חשבנו על שאלות משניות ומהר מאוד הגענו לשלד של מה שאנו רוצים לעשות מבחינה רעיונית. כאשר הגענו לשלב של מימוש השאלות לכדי גרפים התחלנו להבין את גודל הבעייתיות ב-Datasets וכמו שתואר כבר מעלה ביצענו Preprocess אינטנסיבי, בנוסף לכך שהשגנו Dataset נוסף מנתיבי ישראל. לנו היה חשוב שה-Dashboard יהיה מסודר לעין ולכן ניסינו לקבץ כמה שיותר רעיונות תחת אותו plot – בשאלה המשנית הראשונה הוספנו את האופציה לפלטר לפי שנה ובכך אנו נמנעים מיותר מדי פרטים הפרוסים לנגד העיניים. את 2 השאלות המשניות הבאות הכנסנו תחת אותו רעיון של מציאת תבנית כלשהי בכל הקשור לרוחב הכביש ובכך יכולנו להגדיר plot אחד שבעזרת פילטור מתאים עונה על 2 השאלות הנ"ל. בשאלה המשנית האחרונה הוספנו את האופציה עבור המשתמש להשוות בין אזורים שונים בארץ בנוסף לכך שקיימת עבורו האפשרות להתבונן בנתונים בצורה יותר נוחה (Facets). צורת כתיבת הקוד הנ"ל הייתה מאוד מעניינת עבורנו, הרגשנו שאנחנו מבצעים אינטגרציה לרכיבים שונים (אך בעלי קשר כלשהו) של ה-Dashboard תחת אותה קורת גג בעזרת פילטרים.