

איפה ואיך נכון לשכן זוגות צעירים בישובים כפריים בישראל?



קורס: ויזואליזציה של מידע
מגישים: איל ברימן ומאיה ברסלאור
תאריך הגשה: 14/04/2024

תוכן עניינים

3רקע על הנתונים ומטרתם
3רקע
4נתונים
4איסוף הנתונים
4הבנת הנתונים
7הכנת הנתונים
8ויזואליזציות
8ויזואליזציה 1 - התפלגות המדד הסיוציו-אקונומי במחוזות המרכז, הדרום וחיפה צריכים איזון
9ויזואליזציה 2 - גודל האוכלוסייה בגילאי 0-45 מפולגים שונה בין ישובים ערבים ויהודים
12ויזואליזציה 3 - זרימה בין אשכולות, דתות ואזורים
13ויזואליזציה 4 - אחוז הזכאים לבגרות ומדד ג'יני- גדלים עם מספר היישובים במועצה האזורית
14ויזואליזציות אינטרקטיביות
14ויזואליזציה 5 - תרשים מפה- ישובים כפריים כתלות בגודל האוכלוסייה אשכול ודת
ויזואליזציה 6 - מפת חום אינטראקטיבית של אשכול ואזור לפי ממוצע מספר האנשים מקבוצת הגיל
16הנבחרת

רקע על הנתונים ומטרתם

רקע

ישובים קהילתיים מהווים אלטרנטיבה לחיי העיר היקרים והרועשים, ומציעים חיי קהילה וסביבה עוטפת ומשפחתית באווירה כפרית. בחלק מהמועצות האזוריות והמקומיות המרוחקות יש אף אינטרס לאומי לשיכון זוגות צעירים בישובים שבתחום שיפוטן, על מנת להצמיח קהילות צעירות ומתחדשות. בין המקדמות אינטרס זה, נמנות עמותות כגון קק"ל אנגליה, תנועת אור, עמותת קדמה ועמותת איילים, כמו גם ארגונים ממשלתיים, ביניהם משרד הנגב והגליל והחטיבה להתיישבות. גופים אלו עוסקים ביצירת מסגרות שונות לשילוב של צעירים במועצות אזוריות וישובים קהילתיים. זאת, באמצעות הקמת גרעיני התיישבות וכפרי סטודנטים, קיום אירועי תרבות, והצמחת מקומות פנאי והזדמנויות תעסוקה.

ניתן לחלק מועצות אזוריות לפי היכולת התקציבית שלהן, הנובעת מגובה ההכנסות של התושבים בתחומן, כמות ואופי העסקים שקיימים בהן והתעשייה והבסיסים הצבאיים שמספקים לחלקן מקור הכנסה משמעותי. אי לכך, במועצות האזוריות המאגדות כמה ישובים קהילתיים לידי גוף אדמיניסטרטיבי אחד, עשויים להיווצר פערים כלכליים משמעותיים בין ישובים. על מנת לגשר על פערים אלו, בישובים החלשים יותר גובר האינטרס של המועצה לשיכון זוגות צעירים, על מנת לשפר את אופי ומעמד הישוב. אינטרס זה משתקף במענקים והטבות הן מטעם המועצה והן מטעם המדינה. אמנם גם ישובים חזקים יותר מהווים חלופה מתאימה לחיים עירוניים, אך מגורים בהם לרוב לא יכללו קבלת מענקים ותמריצים.

אי לכך, גם למתיישבים בישובים קהילתיים וגם לגופים המיישבים מערכת שיקולים המשפיעה על הבחירה: מחד, זוגות צעירים אשר מתחילים בתהליך החיפוש צריכים לקחת בחשבון גורמים רבים שחלקם תלויים באזור הגיאוגרפי, חלקם באופי הישוב והקהילה וחלקם בשירותי המועצה, בטרם יכנסו לתהליך קבלה מחייב וארוך לאחד היישובים. מאידך, עבור העמותות והארגונים המיישבים נחוצה בהירות בהבנת הגורמים המשפיעים על התכנון האזורי בתחומן, על מנת להכווין מתיישבים באופן המיטבי.

על כן, בחרנו לייצר ויזואליזציות אשר יוכלו לשמש את הארגונים המלווים והמסייעים בתהליך ההתיישבות והקליטה של זוגות צעירים. ארגונים אלו מעוניינים לייצר התאמה מיטבית בין הזוגות לבין היישובים, ולקבל החלטות אופרטיביות בנוגע למוקדי הפעילות שלהם לפי: יישובים, חתכי אוכלוסיה, מועצות אזוריות ומחוזות.

נתונים

איסוף הנתונים

הנתונים בהם השתמשנו בפרויקט נאספו ממספר מקורות שונים שנלקחו מתוך האתר של לשכה המרכזית לסטטיסטיקה ואתר מאגר המידע הממשלתי. הנתונים נאספו בצורה טבלאית בקבצי CSV, XLSX ו-PDF שהומר לקובץ Excel על מנת להשמישו.

הנתונים יוצאו לארבע טבלאות שונות: טבלת תושבים בישראל לפי יישובים וקבוצות גיל, טבלת יישובים, טבלת מדד חברתי כלכלי ואשכולות לפי יישובים וטבלת רשויות מקומיות בישראל. את שלושת הטבלאות הראשונות מיזגנו לטבלה אחת באמצעות כתיבת קוד SQL (שימוש ב-JOIN), לה נקרא טבלת היישובים (המכילה יישובים קהילתיים בלבד). כמו כן, קיימת טבלה נוספת המכילה נתונים על רשויות מקומיות שהשארנו כפי שהיא - אך ניתן לקשר אותה לטבלה הממוזגת לפי הרשות המקומית אליה שייך כל יישוב.

[קובץ PDF של הלמ"ס - מדד חברתי כלכלי ואשכול של יישובים במועצות אזוריות](#)

[מאגרי מידע ממשלתיים - קובץ יישובים 2019](#)

[מאגרי מידע ממשלתיים - תושבים בישראל לפי יישובים וקבוצת גיל](#)

[הרשויות המקומיות בישראל - גיליון של נתונים פיזיים ונתוני אוכלוסייה](#)

[קישור ל-github: https://github.com/mayabres/Visualization-Project.git](https://github.com/mayabres/Visualization-Project.git)

הבנת הנתונים

הנתונים אשר נאספו מציגים מידע אודות יישובים ומועצות אזוריות, על מנת להכווין עמותות המעורבות בתהליכי יישוב וקליטת אנשים ביישובים קהילתיים. באמצעותם, הגופים המיישבים יוכלו לבצע התאמה מיטבית בתהליכי השילוב לתוך האוכלוסיות ביישובים השונים ולמקד את הפעילות שלהם בהתאם לאידיאולוגיה שמקדמים. כל רשומה בכל אחת מבין הטבלאות מייצגת יישוב או מועצה מקומית/אזורית.

טבלת תושבים בישראל לפי יישובים וקבוצות גיל

בטבלה זו רשומה מייצגת יישוב, ומכילה נתונים שונים אודותיו. ביניהם, מספר תושבים לפי קבוצות גיל שונות. טבלה זו מכילה משתנים שונים, אך לצורך היוזאליזציות בחרנו להשתמש במשתנים הנ"ל:

שם שדה	סוג משתנה	טווח ערכים	הסבר
שם_יישוב	String		שם היישוב בעברית
סה"כ	Int		סה"כ אנשים מכל קבוצות הגילאים המפורטות
0_5	Int	0 עד סה"כ	מספר התושבים ביישוב בטווח הגילאים 0 עד 5

מספר התושבים ביישוב בטווח הגילאים 6 עד 18	0 עד סה"כ	Int	6_18
מספר התושבים ביישוב בטווח הגילאים 19 עד 45	0 עד סה"כ	Int	19_45
מספר התושבים ביישוב בטווח הגילאים 46 עד 55	0 עד סה"כ	Int	46_55
מספר התושבים ביישוב בטווח הגילאים 56 עד 64	0 עד סה"כ	Int	56_64
מספר התושבים ביישוב בטווח הגילאים 65 ומעלה	0 עד סה"כ	Int	65_

טבלת יישובים - מדד חברתי כלכלי וחלוקה לאשכולות

בטבלה זו כל רשומה מייצגת יישוב, והיא מכילה את האשכול ואת ערך המדד החברתי-כלכלי של כל יישוב.

שם שדה	סוג משתנה	טווח ערכים	הסבר
שם יישוב	String		שם היישוב בעברית
ערך המדד	Double	בין -3.11 ל- 2.645	מדד חברתי כלכלי של הלמ"ס לפי ישוב- מחושב על בסיס נתונים דמוגרפיים, השכלה וחינוך, תעסוקה, אבטלה וגמלאות. ציון גבוה מעיד על מצב חברתי כלכלי טוב ביישוב בעוד שציון נמוך מעיד על מצב כלכלי חברתי ירוד ביישוב. במועצה אזורית יכולים להיות כמה יישובים עם מדדים חברתיים כלכליים שונים.
אשכול	Int	1-10	סיווג (clustering) על בסיס מדד חברתי-כלכלי. 1- האשכול נמוך ביותר, 10- האשכול הגבוה ביותר.

טבלת יישובים

בטבלה זו כל רשומה מייצגת יישוב או עירייה. בין הנתונים שמציגה, נמנים בין היתר מספר התושבים ביישוב בחלוקה לדתות, דת היישוב, שנת הקמה וגובה ממוצע. מבין המשתנים שטבלה זו מכילה, בחרנו להשתמש בבאים:

שם שדה	סוג משתנה	טווח ערכים	הסבר
שם יישוב	String		שם היישוב בעברית
דת יישוב	Int	1-4	דת היישוב, כאשר: 1: יהדות 2: איסלאם 3: בדואים 4: אחר - מעורב
סך הכל אוכלוסייה 2019	Int		מספר התושבים הכולל ביישוב

טבלת מועצות אזוריות

בטבלה זו נמצאים פריטי מידע רבים על המועצות האזוריות, המועצות המקומיות והעיריות ברחבי הארץ. מתוכם, בחרנו להתמקד בשדות הבאים:

שם שדה	סוג משתנה	טווח ערכים	הסבר
שם הרשות	String		שם הרשות בעברית
מעמד מוניציפלי	String	קטגוריאלי	עירייה/מועצה מקומית/מועצה אזורית
Percentage entitled to a bagrut certificate	Double	0-100	אחוז זכאות לבגרות

Gini index	Double	0-1	מדד המראה חוסר שוויון בהכנסות 0 - שוויון מלא 1 - חוסר שוויון מלא
אשכול	Int	1-10	1 - הנמוך ביותר 10 - הגבוה ביותר

הכנת הנתונים

ביצענו במספר טבלאות עיבוד מקדים לפני איחוד הטבלאות לאחת.

טבלת יישובים

בטבלה זו הסרנו רשומות בהן שם היישוב היה חסר, וכן רשומות בהן הופיעו תוספות שונות לשם היישוב אשר לא היו רלוונטיות (אזור, מ"א, של"ש, *). לאחר הסרות אלו הסרנו את הכפילויות הקיימות בטבלה.

טבלת יישובים - מדד חברתי כלכלי וחלוקה לאשכולות

על מנת לעבוד עם טבלה זו ביצענו המרה מקובץ PDF לקובץ excel.

שמות העמודות אשר הופיעו במסמך ה-PDF.

יישוב בתוך מועצה אזורית LOCALITY IN REGIONAL COUNCIL							מועצה אזורית REGIONAL COUNCIL				
אשכול (מ-1 עד 10) CLUSTER (1 to 10)	ערך מדד INDEX VALUE	אוכלוסייה* POPULATION*	צורת יישוב TYPE OF LOCALITY	NAME OF LOCALITY	שם יישוב	סמל יישוב CODE OF LOCALITY	אשכול (מ-1 עד 10) CLUSTER (1 to 10)	אוכלוסייה* POPULATION*	NAME OF REGIONAL COUNCIL	שם מועצה אזורית	מעמד מוניציפלי MUNICIPAL STATUS

מספר הערות כלליות לגבי המסמך שהופיעו גם הן בקובץ המומר.

* THE POPULATION OF SOCIO-ECONOMIC INDEX 2015 DOES NOT INCLUDE RESIDENTS OF NURSING HOMES, PENITENTIARY INSTITUTIONS, AND SIMILAR.

* אוכלוסיית המדד החברתי-כלכלי 2015 לא כוללת דיירי מוסדות סיעודיים, בתי סוהר וכדומה.

** LOCALITY WITH LESS THAN 120 RESIDENTS. SOCIO-ECONOMIC INDEX 2015 WAS CALCULATED SINCE THE SOCIO-ECONOMIC INDEX 2013 HAS BEEN CALCULATED FOR THIS LOCALITY.

** על אף שאוכלוסיית היישוב מונה פחות מ-120 תושבים, חושב מדד חברתי-כלכלי 2015, כי חושב לו גם מדד חברתי-כלכלי 2013.

.. SOCIO-ECONOMIC INDEX 2015 WAS NOT CALCULATED.

.. לא חושב מדד חברתי-כלכלי 2015.

אפיון יחידות באגרגציות וסיווג לפי הרמה החברתית-כלכלית של האוכלוסייה בשנת 2015

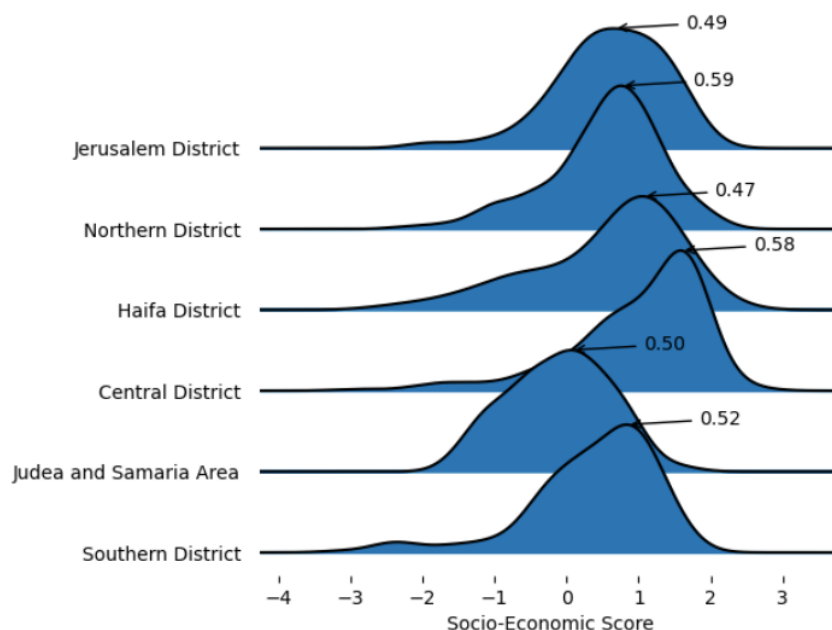
לאחר עיבוד זה שילבנו שלוש טבלאות (יישובים, תושבים, ומדד חברתי כלכלי) לטבלה אחת על מנת שנוכל לייצר את הויזואליזציות. את האיחוד ביצענו באמצעות שאילתת SQL עם שימוש ב-JOIN ולאחר מכן ב-R לטובת ייצוא של הטבלה המאוחדת לקובץ excel.

לאחר האיחוד שמנו לב כי ישוברים רבים הושמטו. הבנו כי ישנם יישובים המאופיינים באופן שונה בין שתי הטבלאות (בעיקר תוספת/החסרת י' או ו') ועל כן היה צריך לערוך את שמות יישובים אלו באופן ידני כך שיתאימו.

לאחר מכן הרצנו על עמודת השמות של היישובים קוד פייתון בעזרת ספריית geopy.geocoders למציאת קואורדינטות ליישובים אשר הוכנסו לעמודות חדשות של קו אורך וקו רוחב. בנוסף מיינו את נתוני הנפות של כל היישובים ויצרנו שדה חדש בשם region עבור כל יישוב, על סמך הנפה: 1- ירושלים, 2- צפון 3- חיפה, 4- מרכז, 5- יהודה ושומרון, 6- דרום (לדוגמה נפת רמאללה- יהודה ושומרון-5, נפת יזרעאל- צפון-2).

ויזואליזציה 1 - התפלגות המדד הסוציו-אקונומי במחוזות המרכז, הדרום וחיפה צריכים איזון

Balancing Socio-Economic Disparities in the South Region, Haifa Region, and Central Region



WHAT?	התפלגות מדד חברתי כלכלי ביישובים כפריים (אינטרוולי) לפי מחוזות (קטגוריאלי).
WHY?	הצגת הבדלי ההתפלגות במדד החברתי כלכלי של היישובים לפי מחוזות.
HOW?	Reduce - אגריגציית מדד היישובים לפי מחוז Encode - קידוד של הנתונים וסידורם כדי להציג את התפלגות המחוזות (אזורים) לפי מדד סוציו-אקונומי.

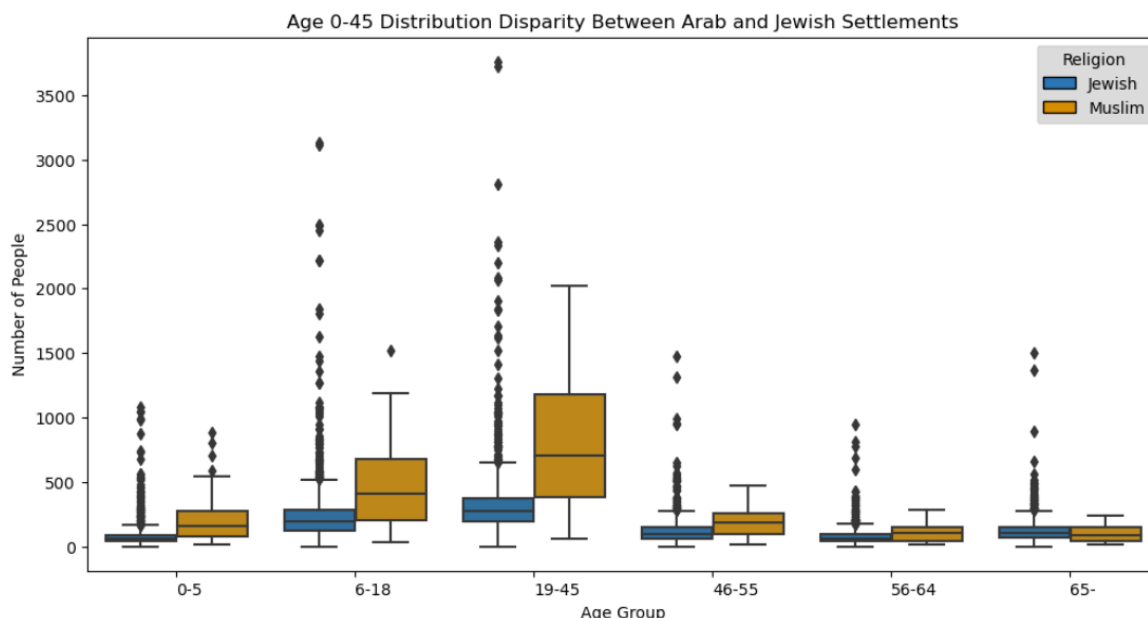
ויזואליזציה זו מתארת את התפלגות המדד הכלכלי-חברתי של מרבית היישובים הקהילתיים בארץ (הן של מועצות מקומיות והן של ישובים במועצות אזוריות), בחלוקה לפי מחוזות: ירושלים, צפון, חיפה, מרכז, יהודה ושומרון והדרום. מדובר בתרשים מסוג התפלגות, שנועד לשרת את העמותות והארגונים המיישבים ואת מקבלי ההחלטות. מטרת התרשים היא להציג באילו אזורים כדאי למקד הקמת גרעיני התיישבות על מנת לחזק את המחוז. סביר שהמסקנה תוביל למיקוד במחוזות המאופיינים בהתפלגות בעלת זנב שמאל דומיננטי כמו אזור מרכז, חיפה, והדרום.

צורת התפלגות שאינה פעמונית מעידה על חוסר איזון בין היישובים השונים במחוז מבחינת המדד, ועבור מקבלי ההחלטות, מדובר בהזדמנות "לאזן" את ההתפלגות המחוזית.

כמו כן, ניתן לקבל מהתרשים פרופיל רצוי של אוכלוסיות להצבה מתאימה לפיתוח וחיזוק אזורי. כך, יתכן כי אנשים משכילים בעלי הכנסה גבוהה יתרמו יותר במחוזות המתאפיינים בפערים גדולים וחוסר איזון בין היישובים. (למשל, מחוז דרום: ביר הדאג' - מדד נמוך מאוד, מיתר - מדד גבוה מאוד).

יצירת התרשים התבצעה באמצעות חבילת jupyter בפיתוח וקדמה לכך פקודת אקסל אשר מיפתה עבור כל נפה בה נמצא היישוב את האזור המתאים. בנוסף הוספנו את הערך המקסימלי של כל התפלגות בתוספת חץ המצביע על ערך המקסימום, על מנת שיהיה ניתן לערוך השוואה בין ההתפלגויות גם על בסיס גובה.

ויזואליזציה 2 - גודל האוכלוסייה בגילאי 0-45 מפולגים שונה בין ישובים ערבים ויהודים

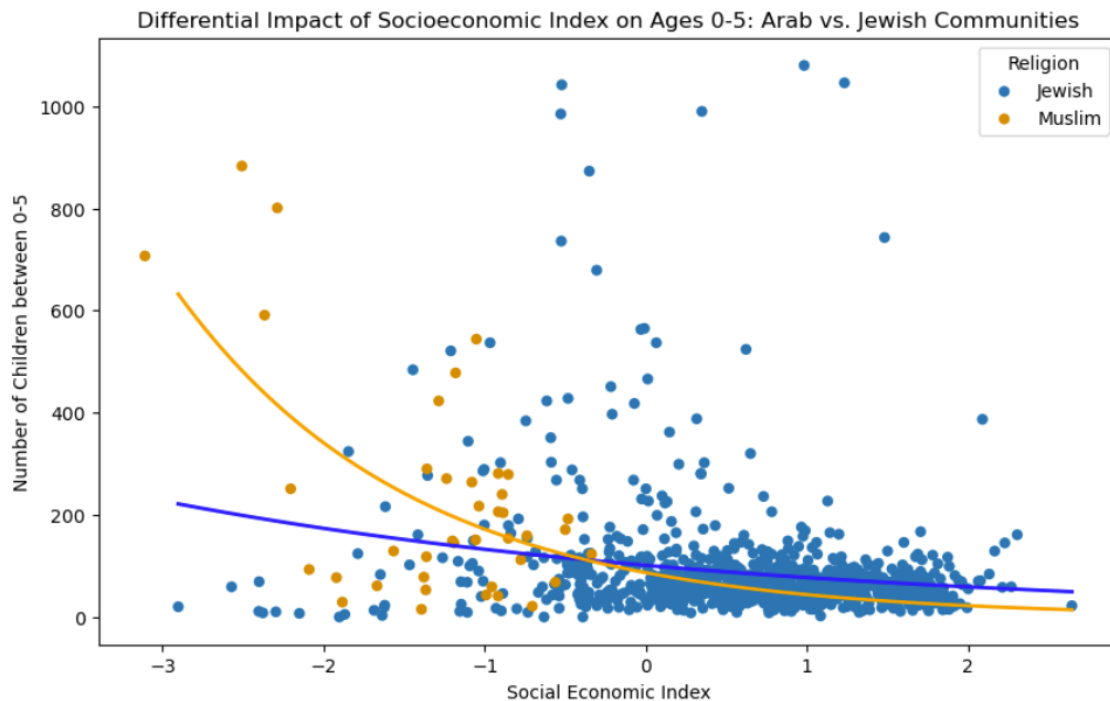


WHAT?	התפלגות מספר אנשים בקבוצות גילאים שונות (יחס) בישובים כפריים לפי דת (קטגוריאלי).
WHY?	הצגת הבדלי התפלגות בין קבוצות הגיל השונות בין יהודים ומוסלמים.
HOW?	סינון הנתונים, והצגת צורת הקופסה בצבע שונה לכל דת.

ויזואליזציה זו משלימה את המפה האינטרקטיבית ומתארת את ההתפלגות של מספר האנשים בקבוצות גיל שונות בישובים כפריים ערביים אל מול ישובים יהודיים. מטרת התרשים היא ציור תמונת מצב עבור מקבלי ההחלטות ממשרד הנגב והגליל והחטיבה להתיישבות אודות השוני בהתפלגויות בין ישובים כפריים ערבים ובין ישובים כפריים יהודים, שניכר בפרט בקבוצות הגיל 0-45. פער זה מתאפיין בעיקר בכמות הגדולה של ילדים בגילאי 0-5 ו-6-18 ביישובים הערביים אל מול הישובים היהודים, בניכוי ישובים יהודיים חריגים (התנחלויות וישובים קהילתיים גדולים). פער נוסף טמון בשונות באופי הישובים; בעוד קבוצת הישובים היהודיים מאופיינת בהטרוגניות, כך שקיימים יישובים רבים מסוגים שונים, ניתן לראות כי ביישובים הערביים ישנן תצפיות חריגות מועטות, והם מהווים קבוצה הומוגנית יותר. כמו כן, ניתן לראות במפה האינטרקטיבית שישובים ערביים מתאפיינים באשכול נמוך ועל כן גם במדד חברתי-כלכלי נמוך.

לכן, רצינו לבדוק האם קיים הבדל בקשר שבין מדד חברתי כלכלי וריבוי ילדים בין ישובים ערביים ויהודיים, ועל כן ביצענו את שתי הוויזואליזציות הבאות. הוחלט לבצע תרשים רגרסיה (GLM פואסון- שכן מדובר בנתונים חיוביים המתארים ספירה) על גבי תרשים פיזור של מספר התושבים בישוב בני 0-5 ו-6-18 כתלות בממד החברתי כלכלי.

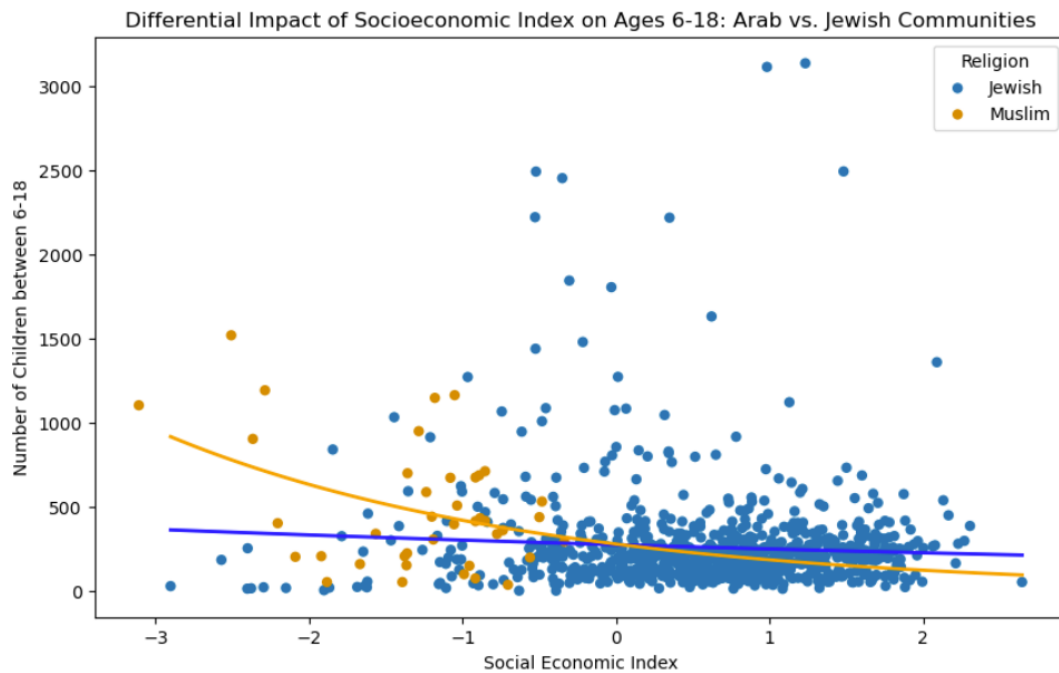
גודל האוכלוסיה בגילאי 0-5 מושפע יותר מהמדד הסוציו אקונומי בישובים הערבים מאשר ישובים יהודים:



WHAT?	קשר ומגמה בין מספר אנשים בקבוצות גיל 0-5 (יחס) בישובים כפריים כתלות בממד הסוציו-אקונומי (אינטרוולי), לפי דת (קטגוריאלי).
WHY?	השוואת הקשר בין מספר ילדים בגיל 0-5 לבין מדד סוציו-אקונומי בין ישובים יהודים וישובים ערבים.
HOW?	Encode - קידוד וארגון של הנתונים, צביעתם לפי דת ואגריגציית הנתונים (נקודות) לידי קו רגרסיה המראה קשר.

גודל האוכלוסיה בגילאי 18-6 מושפע יותר מהמדד הסוציו-אקונומי בישובים הערבים מאשר ישובים

יהודים:

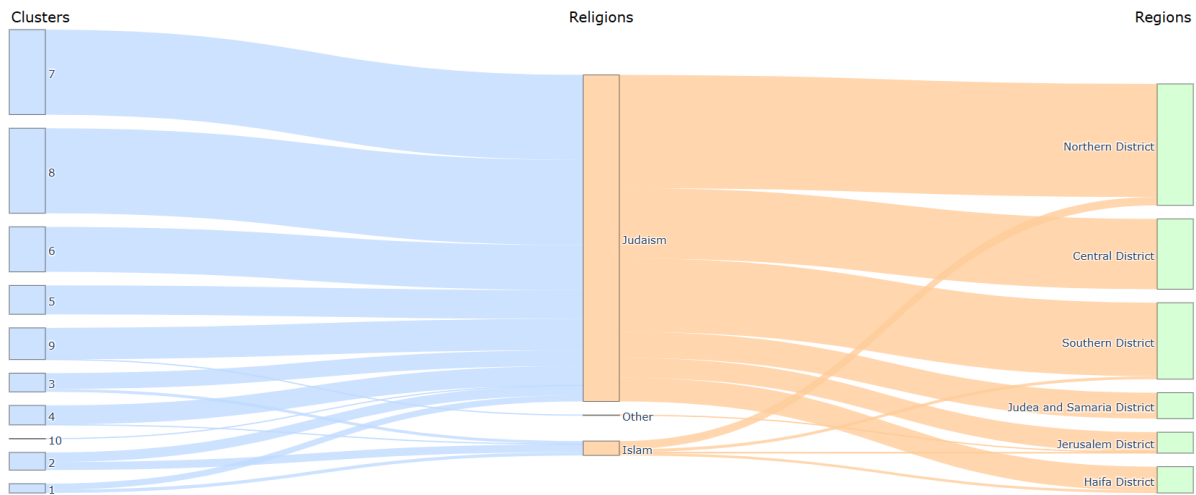


WHAT?	קשר ומגמה בין מספר אנשים בקבוצות גיל 18-6 (יחס) בישובים כפריים כתלות במדד הסוציו-אקונומי (אינטרוולי), לפי דת (קטגוריאלי).
WHY?	השוואת הקשר בין מספר ילדים בגיל 18-6 לבין מדד סוציו-אקונומי בין ישובים יהודים וישובים ערבים.
HOW?	Encode - קידוד וארגון של הנתונים, צביעתם לפי דת ואגריגציית הנתונים (נקודות) לידי קו רגרסיה המראה קשר.

המסקנה משלושת גרפים אלו היא הצורך בשילוב אוכלוסיה ערבית צעירה, משכילה יותר ומסורתית פחות (על מנת שיהיו פחות ילדים בכל משפחה), בישובים הערבים הכפריים או לשלב אוכלוסיה זו בהקמת ישובים כפריים חדשים בנוסף לאלו הוותיקים אשר יסייעו בהעלאת המדד וגם יאזנו את התפלגות מספר הילדים. זאת, על מנת לאפשר התיישבות ערבית כפרית בעלת מדד חברתי-כלכלי גבוה יותר.

ויזואליזציה 3 - זרימה בין אשכולות, דתות ואזורים

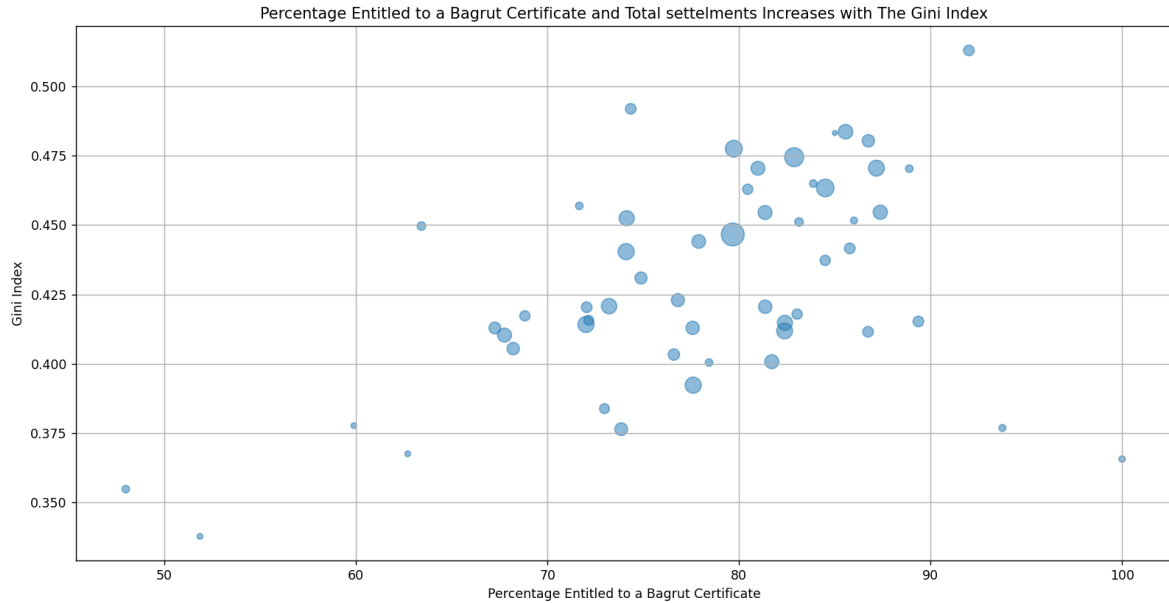
Flow between Clusters, Religions and Regions



ויזואליזציה זו מתארת את הקשר בין אשכול היישוב לדת היישוב ובין הדת לאזור היישוב, וכן את כמות היישובים הנמצאים בכל דת/אשכול/אזור. בחרנו להציג זאת באמצעות תרשים sanky כדי לראות את זרימת הנתונים בין שלושת המאפיינים האלו, את הקשרים השונים ביניהם ואת כמות ופרופורציית היישובים בכל אחת מהקטגוריות המתוארות בו. ניתן להסיק מתרשים זה כי יישובים מוסלמים מאופיינים באשכול נמוך ואינם ממוקמים באזור המרכז ובאזור יהודה ושומרון. כמו כן, ניתן לראות כי יישובים רבים בעלי אשכול גבוה (מ-7 ומעלה), אלו יישובים יהודים. בנוסף, נראה כי רוב היישובים הכפריים הנמצאים באזור הצפון ובאזור הדרום. מסקנות אלו יסייעו לעמותות והארגונים הממשלתיים בקבלת החלטות אודות מיקוד הפעילות שלהם.

WHAT?	זרימת מספר יישובים לפי אשכול, דת ואזור. חלוקה של היישובים לפי אשכולות (אורדינאלי), לפי דתות (קטגוריאלי) ולפי מחוזות (קטגוריאלי).
WHY?	נרצה לבצע השוואה בין זרימות מספר היישובים בתרשים בין דת היישוב לאשכול, ובין דת היישוב לאזור על מנת לקבל תמונת מצב אודות האזורים השונים כדי לאפשר קבלת החלטות שקולות יותר של העמותות.
HOW?	Encode - קידוד על ידי ארגון הנתונים. נציג זאת באמצעות תרשים Sanky, אשר מציג זרימה של נתונים. תרשים זה מכיל עמודת מקור (Clusters) ועמודת יעד (Regions). כמו כן, נציין כי הצבעים מציינים את המשתנים (ולא את הערכים). תרשים זה גם אינטראקטיבי ועומד בפני עצמו. בנוסף, באמצעות מעבר עכבר על ה-flows ועל הקטגוריות השונות ניתן לראות באופן ממוקד יותר את זרימת הנתונים בין אשכול, דת ואזור.

ויזואליזציה 4 - אחוז הזכאים לבגרות ומדד ג'יני- גדלים עם מספר היישובים במועצה האזורית



הסבר מקדים - מדד ג'יני: מדד ג'יני הוא מדד לאי שוויון בחלוקת הכנסות. מדד זה הוא בין 0 ל-1: כאשר ערכו 1 קיים אי שוויון מלא, וכאשר ערכו 0 קיים שוויון מוחלט. מדד זה מחושב לפי היחס בין השטח הכלוא בין עקומת לורנץ לקו 45 מעלות, ושטח המשולש כולו (זה שמתחת לקו 45 מעלות ושטחו הוא 0.5).

ויזואליזציה זו מספקת לעמותות תמונת מצב כללית על המועצות האזוריות לגבי הקשר בין אחוז הזכאים לבגרות, מספר היישובים במועצה ומדד ג'יני. ניתן לראות בויזואליזציה זו כי ככל שגדל אחוז הזכאים לבגרות, מספר היישובים גדול יותר וערכו של מדד ג'יני גדל (קיים אי שוויון גדול יותר במועצה). זוגות צעירים עשויים להעדיף דווקא לגור במועצה אזורית קטנה - בה ככל הנראה הכל קרוב לבית ומכירים את רוב אוכלוסיית המועצה. בנוסף, זוגות מחפשים מקום מגורים להשתקע בו ולהקים משפחה יחפשו ככל הנראה מקום עם בתי ספר טובים - אחוז גבוה של זכאים לבגרות במועצה מסוימת עשוי להעיד על איכות בתי הספר האזוריים וכן על חוזק האוכלוסייה.

WHAT?	קשר בין אחוז זכאים לבגרות במועצה אזורית (יחס), מדד ג'יני (אינטרוולי) ומספר יישובים במועצה אזורית (אינטרוולי).
WHY?	באמצעות ויזואליזציה זו רצינו לבחון את השערותנו שקיים קשר בין שלושת מאפיינים אלו במועצות אזוריות. משתנים אלו רלוונטיים בבחירת מקום מגורים עבור זוגות צעירים אשר מעוניינים למצוא מקום להשתקע בו ולהקים משפחה. חלק מהסיבות לבחירת מקום מגורים יכולות להיות אחוז הזכאים לבגרות (אשר עשוי להעיד על רמת החינוך במועצה אזורית זו), מדד ג'יני אשר מעיד את רמת חוסר השוויון בהכנסות ומספר היישובים במועצה (מעיד על גודל המועצה).
HOW?	Encode - קידוד וארגון של הנתונים לצורך קבלת תרשים פיזור.

	Size - גודל כל נקודה (תצפית - מועצה אזורית) הוא לפי כמות היישובים במועצה.
--	---

ויזואליזציות אינטרקטיביות

ויזואליזציה 5 - תרשים מפה- ישובים כפריים כתלות בגודל האוכלוסייה אשכול ודת

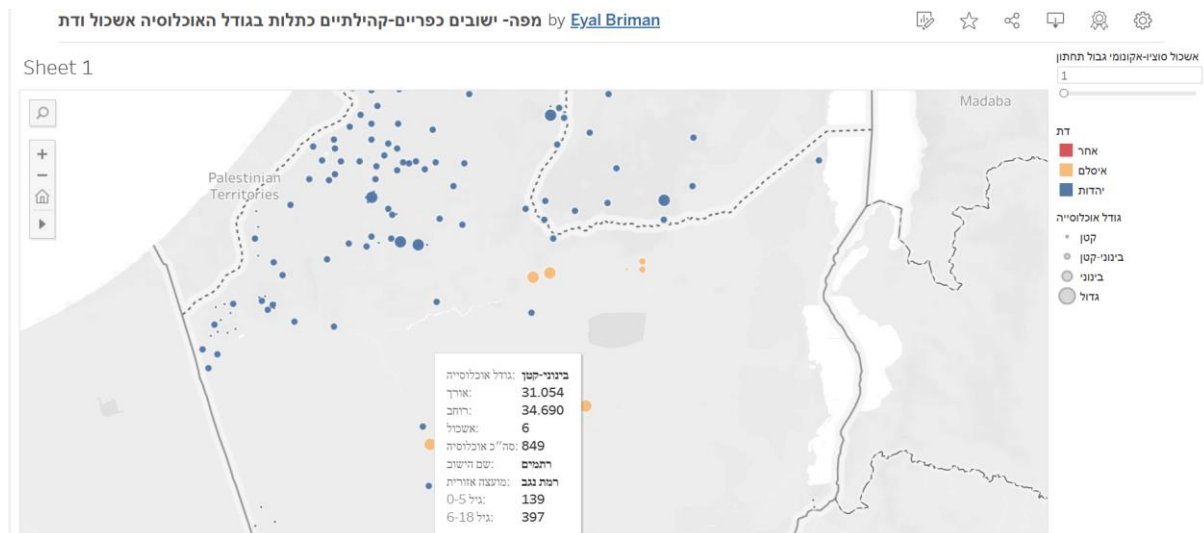
קישור: [מפה- ישובים כפריים-קהילתיים כתלות בגודל האוכלוסייה אשכול ודת | Tableau Public](#) (קובץ Tableau נמצא בGIT) לצורך יצירת תרשים המפה האינטרקטיבי נעשה שימוש בטבלת הישובים (המורכבת כפי שהוסבר קודם, ממספר מקורות), והרצנו על עמודת השמות של הישובים קוד פייתון בעזרת ספריית geopy.geocoders למציאת קואורדינטות לישובים אשר הוכנסו לעמודות חדשות של קו אורך וקו רוחב. לאחר מכן ייצאנו את הטבלה בצירוף שתי העמודות החדשות לאקסל וייבאנו את הטבלה החדשה אל סביבת Tableau שם דאגנו להגדיר הן את השמות של הישובים והן את הקואורדינטות שלהם כ- Dimensions ויצרנו מפה על בסיס הקואורדינטות. לאחר מכן ביצענו תיקוני חריגים באופן ידני על ידי חיפוש הקואורדינטות של הישוב ב Google Maps ועדכנו את בטבלת אקסל המלאה. לאחר מכן, הפכנו את שדה סה"כ אוכלוסייה ל- Dimension גם כן, ויצרנו שדה מחושב של גודל ישוב לפי הקוד הבא:

גודל אוכלוסייה

```
IF [סהכ] >= 0 AND [סהכ] <= 500 THEN "קטן"  
ELSEIF [סהכ] > 500 AND [סהכ] <= 1500 THEN "בינוני-קטן"  
ELSEIF [סהכ] > 1500 AND [סהכ] <= 5000 THEN "בינוני"  
ELSE "גדול"  
END
```

חלוקה זו התבצעה באופן שנראה לנו הגיוני לפי טווח הנתונים הקיים (בין 100 אנשים בישוב הקטן ביותר ל 12000 בישוב הגדול ביותר).

את גודל היישוב (שדה מחושב) הוספנו ל- Marks לפי גודל (tool-tips) והוספנו מקרא. עוד הוספנו ל Marks דת לפי צבע והקפדנו להתאים את הצבעים לפי דת באופן התואם את הגרפים הלא אינטרקטיבים שהצגנו קודם. הגדרנו את אוכלוסיית הילדים בגיל 0-5 ובגיל 6-18 כ Dimensions שכן הן האוכלוסיות המעניינות את קהל היעד שלנו (העמותות והארגונים המיישבים), שמטרתם ליישב זוגות צעירים (שסביר כי יחפשו קהילה צעירה- והורים לילדים), והוספנו אותם ל Marks. הוספנו גם את המועצה האזורית, סה"כ האוכלוסייה בישוב ואת האשכול ל Marks ויצרנו פילטר על האשכול (גבול תחתון- ישובים הקטנים ממנו יהיו מוסתרים) באמצעות סליידר. את המפה העלנו ל Tableau Server. להלן מצורפת דוגמא לתפעול כלי הויזואליזציה.

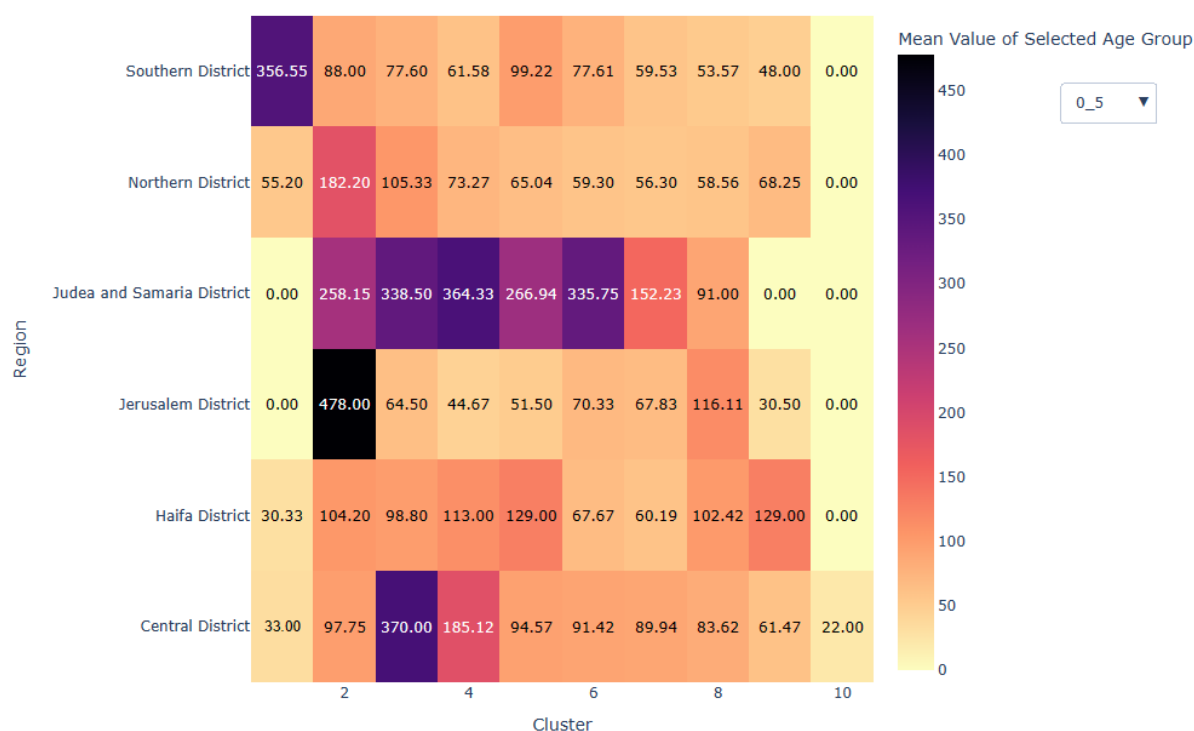


WHAT?	הצגת ישובים לפי מיקומם במפה לפי גודל אוכלוסיה (אינטרוולי) ובסיכון לפי אשכול (אורדינאלי) ומאפיינים נוספים אותם ניתן לראות באמצעות עמידה עם העכבר על היישוב הרצוי.
WHY?	ארגון נתונים נבחרים אודות היישובים לטובת קבלת החלטות של זוג או עמותה בקונטקסט גיאוגרפי.
HOW?	<p>Encode - שימוש בנתונים כפי שהם והצגתם על מפה.</p> <p>Manipulate - יצירת מאפיין חדש (גודל היישוב) על סמך מספר התושבים בו.</p> <p>Map:</p> <p>Color - צביעה של היישובים לפי מאפיין הדת - יהדות נצבע בכחול, איסלם נצבע בכחול ואחר (מעורב) נצבע באדום.</p> <p>Size - גודל היישוב נקבע לפי המאפיין שיצרנו עבור ויזואליזציה זו - גודל היישוב.</p> <p>באמצעות הצגת נתוני היישובים על מפה, אנו מאפשרים לעמותות השונות לבחון כל יישוב בנפרד תוך הסתכלות על היישובים שסביבו וכך לבחון את התאמתו למגורי זוג צעיר אשר מחפש את ביתו החדש. ניתן לראות את נתוני היישוב הרצוי על ידי מעבר עכבר על היישוב הרצוי - במעבר עכבר על אותו יישוב תיפתח תיבת טקסט עם הפרטים השונים על אותו יישוב. בנוסף, איך ביצענו זאת נכתב באופן מפורט מעל טבלה זו.</p>

ויזואליזציה 6 - מפת חום אינטראקטיבית של אשכול ואזור לפי ממוצע מספר האנשים מקבוצת הגיל

הנבחרת

Interactive Heatmap of Region VS Cluster According to Selected Age Group



זוגות אשר מחפשים מקום מגורים ככל הנראה יבססו את החלטתם על בסיס אזור במדינה בו מעוניינים לגור. יתר על כן, אנשים יעדיפו כפי הנראה להתגורר במקום בעל מעמד סוציו-אקונומי טוב או במקום בו יינתנו תמריצים. בנוסף, זוגות צעירים ירצו ככל הנראה להתגורר במקום בו יש אוכלוסייה צעירה הקרובה לגילם וגם ילדים - במידה ומעוניינים להקים משפחה. לכן, על ידי שילוב שלושת מאפיינים אלו, הזוגות הצעירים והעמותות השונות לקבל תמונה ברורה יותר.

הערה: ויזואליזציה זו נגישה ב-github- קישור נמצא בתחילת העבודה.

WHAT?	הצגת ממוצע מספר האנשים בקבוצת גיל מסוימת, נבחרת, (אינטרוול) אשר גרים ביישובים באזור מסוים (קטגוריאלי) באשכול מסוים (אורדינלי)
WHY?	סידור נתוני אשכולות, אזורים וקבוצות גיל לצורך קבלת תמונת מצב כללית שתסייע לעמותה ולזוג צעיר למקד את החלטת מקום מגורי הזוג.
HOW?	<p>Reduce - אגריגציה של הנתונים ופילטור, על מנת שנוכל להציג בכל פעם את ממוצע מספר האנשים בקבוצת גיל נבחרת, הגרים ביישובים באזור מסוים ובאשכול מסוים.</p> <p>Map:</p> <p>Color - צביעה של ערכי ממוצע מספר האנשים לפי סקאלת צבעים - ממוצע קטן של מספר אנשים בקבוצת הגיל הנבחרת נצבע בצהוב וממוצע גדול נצבע בשחור. ערך הטקסט על כל משבצת נצבע בשחור עבור משבצות עם רקע בהיר ונצבע בלבן עבור משבצות עם רקע כהה.</p>