

## מתודולוגיה של הסקר החברתי

### אוכלוסיית הסקר

כל האוכלוסייה הקבועה של מדינת ישראל, בגיל 20 ומעלה. אוכלוסיית הסקר כוללת גם דיירי מוסדות השוהים במוסד שלא לצרכים טיפוליים (כגון מעונות סטודנטים, מרכזי קליטה ותכניות דיור מוגן לאוכלוסייה המבוגרת). עולים חדשים נכללו באוכלוסיית הסקר, רק אם שהו בארץ לפחות 6 חודשים. החל משנת 2004, אוכלוסיית הסקר כוללת את תושבי מזרח ירושלים. לא נכללים באוכלוסיית הסקר: דיירים במוסדות טיפוליים (כגון בתי אבות, בתי חולים לחולים במחלות כרוניות ובתי סוהר), תושבים ישראלים ששהו בחו"ל למעלה משנה ברציפות בעת הסקר, דיפלומטים ועולים אשר הגיעו ארצה חצי שנה לפני מועד הריאיון לכל המאוחר, שבטי בדווים ותושבים הגרים מחוץ ליישובים.

### שיטת הדגימה

#### סקר חברתי 2002

מרשם התושבים כולל רשומות רבות שאינן שייכות לאוכלוסיית הסקר החברתי (כיסוי עודף), כמו למשל ישראלים החיים באופן קבוע בחו"ל (גם אם מפעם לפעם הם מבקרים לתקופות קצרות בארץ) וכן חלק מתושבי המוסדות ותושבי יישובים מסוימים. כדי לייעל את אומדני הסקר ואת תהליך איסוף הנתונים, זוהו ואתרו חלק מהרשומות העודפות לפני הוצאת המדגם, ונכחו מקובץ מרשם התושבים ששימש כמסגרת לדגימה. הדגימה התבצעה בנפרד עבור יישובים גדולים ויישובים קטנים. יישובים גדולים הוגדרו כיישובים המונים כ-4,700 תושבים בני 20 ומעלה. ביישובים אלה מתגוררים מעל 85% מאוכלוסיית הסקר. מדגם הנפשות בקבוצת היישובים הגדולים הוצא ב-56 שכבות דגימה שהוגדרו לפי שילוב של 2 קבוצות מין (גברים ונשים), 4 קבוצות אוכלוסייה (ערבים, עולים עד 1989, עולי 1990 ואילך ויהודים ילידי ישראל) ו-7 קבוצות גיל (20-24, 25-34, 35-44, 45-54, 55-64, 65-74, +75). לכל שכבה הוקצה מדגם שהוא בקירוב יחסית לגודל השכבה. שכבות דגימה שהוערך שתהיינה חשופות בסקר לשיעורי אי-השבה גבוהים, תוגברו בהתאם. בכל שכבה הוצא מדגם מקרי-שיטתי של נפשות. הדגימה ביישובים קטנים הייתה דו-שלבית, בהתאם לאילוצים תפעוליים. בשלב הראשון נדגמו יישובים בשכבות שהוגדרו לפי מאפיינים גאוגרפיים וצורת יישוב. הקצאת המדגם בין השכבות הייתה יחסית לסך כל בני 20 ומעלה בשכבה. בכל שכבה נדגמו יישובים יחסית לסך כל בני 20 ומעלה ביישוב. בשלב השני נדגמו כ-10 נפשות בכל אחד מהיישובים שנדגמו בשלב הראשון. מאחר שקובץ מרשם התושבים היה מעודכן לינואר 2001 ואוכלוסיית הסקר כללה גם את עולי 2001, הוצא מתוך קובץ עולי 2001 מדגם תוספות של עולים. בסך הכל המדגם כלל 9,518 נפשות במטרה לקבל כ-7,000 משיבים.

## שינויים בשיטת הדגימה בסקר חברתי 2003

המדגם הוצא בדומה לשיטות הדגימה של הסקר החברתי 2002. השינויים העיקריים היו: מסגרת הדגימה התבססה על קובץ מרשם התושבים, המעודכן לתקופה מאוחרת יותר - אפריל 2002. מספר הנפשות המינימלי ביישוב גדל מ-10 נפשות ל-15 נפשות. כתוצאה מכך, יישובים גדולים שבהם מוצא מדגם שכבות חד-שלבי הם אלו עם 7,100 או יותר בני 20 ומעלה (כ-83% מאוכלוסיית הסקר מתגוררים ביישובים אלו).

## שיטת הדגימה בסקר חברתי 2004

גודל המדגם תוכנן כך שמספר המשיבים הצפוי יהיה כ-7,200 בני 20 ומעלה. בשלב ראשון הוגדרו קבוצות תכנון ע"י שילוב של 3 משתנים דמוגרפיים: 5 קבוצות אוכלוסייה (ערבים ממזרח ירושלים, ערבים שלא ממזרח ירושלים, עולי 1990 ואילך, עולים עד 1989 ויהודים ילידי ישראל), 7 קבוצות גיל (20-24, 25-34, 35-44, 45-54, 55-64, 65-74, +75) ו-2 קבוצות מין (גברים ונשים) – בסך הכל 70 קבוצות תכנון. גודל מדגם המשיבים הצפוי לקבוצה תוכנן להיות יחסי לגודל הקבוצה (מספר הנפשות במסגרת בקבוצת תכנון), ותחת האילוץ של מספר משיבים צפוי - 7,200. על סמך ההנחות לשיעורי אי-השבה (נפל) ואי-שייכות לאוכלוסיית הסקר (אפס), בכל קבוצה התקבל הגודל הדרוש לדגימה. בחלק מקבוצות התכנון תוגבר גודל המדגם להורים לילדים בגיל 0-13, על סמך הידוע ממרשם התושבים, זאת כדי להגדיל את מספר המיועדים להשיב לשאלון המתחלף בנושא סידורים לילדים בגיל 0-13. התגבור נעשה על חשבון אלה שאינם הורים מקבוצות תכנון שונות, כדי להבטיח גודל מדגם כמתוכנן. יש להעיר כי תגבור הורי הילדים בגיל 0-13 לא נעשה באופן גורף, כדי לא לפגוע באמידה של תכונות שאלון הגרעין המתייחס לכלל בני 20 ומעלה השייכים להגדרת אוכלוסיית הסקר. בסופו של דבר, התגבור הביא את שיעור ההורים לילדים בגיל 0-13 במדגם לכדי 34% מכלל המדגם, לעומת 29% אילו לא נערך תגבור כלל. הסתברות הדגימה הסופית לנפש משתנה בין קבוצות התכנון ולעתים גם בתוך קבוצות התכנון והיא מבטאת הערכות מוקדמות לאחוזי אי-השבה בקבוצות השונות ומשפיעה על החלטות לגבי תגבור ההורים. הסתברות הדגימה הממוצעת הייתה 1:485, הסתברות הדגימה המקסימלית הייתה 1:305, והמינימלית - 1:920.

נפשות ביישובים שבהם צפויים להתקבל לפחות 15 מקרים במדגם (כלומר עם כ-7,400 בני 20 ומעלה), נדגמו במדגם שכבות חד-שלבי, כאשר כל שכבת דגימה היא קבוצת תכנון. בכל שכבה הוצא מדגם מקרי-שיטתי של נפשות לאחר מיון הקובץ לפי מאפיינים גאוגרפיים. בחלק משכבות הדגימה נקבעה הסתברות הדגימה לנפש בהתאם למידע מהמרשם - אם הורה לילד בגיל 0-13 או לא - וביטאה את ההחלטות בדבר תגבור ההורים. בחלק אחר של שכבות הדגימה, שבהן לא תוגברו הורים ולא הוקטנה ההקצאה כדי לתגבר הורים בשכבות אחרות, הסתברות הדגימה של הנפשות בשכבה הייתה אחידה. בסך הכל כ-83% מהמדגם הוצאו בשיטת דגימה חד-שלבית.

כדי לענות על האילוצים התפעוליים, הוצא מדגם דו-שלבי עבור האוכלוסייה ביישובים המונים פחות מ-7,400 בני 20 ומעלה. בשלב ראשון חולקו היישובים לשכבות גאוגרפיות והוצא מדגם של יישובים יחסית לגודלם. בשלב שני הוצא מדגם מקרי-שיטתי של נפשות ביישובים שנדגמו, לאחר מיון הרשומות לפי מאפייני 70 התכונות לתכנון. הסתברות הדגימה של השלב השני הותאמה כך

שהסתברות הדגימה הסופית (כפל ההסתברויות בשני השלבים) תהיה קבועה. בשיטה זו התקבל גודל מדגם של כ-15 נפשות כנדרש.

בתחילת 2004 הוצא מדגם תוספות של עולים חדשים, אשר עלו לארץ החל מאפריל 2003 ולא נכללו במסגרת הדגימה העיקרית. סה"כ כלל המדגם 9,528 נפשות במטרה לקבל כ-7,200 משיבים לסקר.

### **שיטת הדגימה בסקר חברתי 2005**

המדגם הוצא בשיטה דומה לזו ששימשה בסקר החברתי 2003, אך עם 70 קבוצות תכנון (לפי אותו פירוט שיושם בסקר חברתי 2004). מסגרת הדגימה התבססה על מרשם התושבים המעודכן לאפריל 2004. המדגם תוכנן כדי לקבל כ-7,200 משיבים בני 20 ומעלה. בסך הכל נדגמו לסקר 9,515 בני 20 ומעלה.

### **שיטת הדגימה בסקר חברתי 2006**

המדגם התבסס על 70 קבוצות תכנון, הזרות לאלו של הסקר החברתי 2005. עבור הנושא המתחלף לסקר 2006 – "מטפלים ללא תמורה כספית" התקבל מידע על מטופלים מתוך קובץ נתוני יסוד של משרד הרווחה והשירותים החברתיים, ולמטופלים אלה זוהו קרובי משפחה במרשם, אשר עשויים להוות "מטפלים פוטנציאליים". בכל קבוצת תכנון, "מטפלים פוטנציאליים" קיבלו תגבור על חשבון האחרים, זאת כדי לייעל את אומדני הנושא המתחלף לשנת 2006, אך תחת המגבלה של תכנון המדגם עבור 7,500 משיבים צפויים. בסך הכל נדגמו לסקר 9,499 בני 20 ומעלה.

### **שיטת הדגימה בסקר חברתי 2007**

המדגם הוצא בשיטה זהה לזו של הסקר החברתי 2005, ללא תגבור עקב נושא מתחלף. מסגרת הדגימה התבססה על מרשם התושבים המעודכן לאפריל 2006. המדגם תוכנן על בסיס 70 קבוצות תכנון כדי לקבל כ-7,500 משיבים. בסך הכל נדגמו לסקר 9,391 בני 20 ומעלה.

### **שיטת הדגימה בסקר חברתי 2008**

המדגם הוצא בשיטה זהה לזו של הסקר החברתי 2007, ללא תגבור עקב נושא מתחלף. מסגרת הדגימה התבססה על מרשם התושבים המעודכן לאפריל 2007. המדגם תוכנן על בסיס 70 קבוצות תכנון כדי לקבל כ-7,500 משיבים. בסך הכל נדגמו לסקר 9,389 בני 20 ומעלה.

### **שיטת הדגימה בסקר חברתי 2009**

המדגם הוצא בשיטה זהה לזו של הסקר החברתי 2008 ללא תגבור עקב נושא מתחלף. מסגרת הדגימה התבססה על מרשם התושבים המעודכן לאפריל 2008. המדגם תוכנן על בסיס 70 קבוצות תכנון כדי לקבל כ-7,500 משיבים. בסך הכל נדגמו לסקר 9,340 בני 20 ומעלה. בדיעבד, התברר כי בשנת 2009 נדגמו מספר גדול יחסית של יישובים לא-עירוניים יהודיים עם אוריינטציה דתית, שהביאו לאומדן גבוה יחסית של פרטים שהגדירו עצמם כדתיים בשאלון הסקר, לעומת שנת 2008.

## שיטת הדגימה בסקר חברתי 2010

גודל המדגם תוכנן כך שמספר המשיבים הצפוי יהיה כ-7,500 בני 20 ומעלה. בשלב ראשון הוגדרו קבוצות תכנון ע"י שילוב של 3 משתנים דמוגרפיים: 5 קבוצות אוכלוסייה (ערבים ממזרח ירושלים, ערבים שלא ממזרח ירושלים, עולים עד 1989, עולי 1990 והלאה ויהודים ילידי ישראל), 7 קבוצות גיל (20-24, 25-34, 35-44, 45-54, 55-64, 65-74, +75) ו-2 קבוצות מין (גברים ונשים) – סך הכול 70 קבוצות תכנון. גודל מדגם המשיבים הצפוי לקבוצה תוכנן להיות יחסי לגודל הקבוצה (מספר הנפשות במסגרת בקבוצת תכנון), ותחת האילוץ של מספר משיבים צפוי - 7,500. על סמך ההנחות לשיעורי אי-השבה (נפל) ואי-שייכות לאוכלוסיית הסקר (אפס), בכל קבוצה התקבל הגודל הדרוש לדגימה. הסתברות הדגימה הממוצעת הייתה 1:531, הסתברות הדגימה המקסימלית הייתה 1:380 והמינימלית - 1:602.

הנפשות ביישובים שבהם היו צפויים להתקבל לפחות 15 מקרים במדגם (כלומר שבהם כ-8,000 נפשות בני 20 ומעלה), נדגמו במדגם שכבות חד-שלבי, כאשר כל שכבת דגימה היא קבוצת תכנון. בכל שכבה הוצא מדגם מקרי-שיטתי של נפשות לאחר מיון הקובץ על ידי מאפיינים גאוגרפיים. בסך הכל, כ-84% מהמדגם הוצא בשיטה דגימה חד-שלבית.

כדי לענות על האילוצים התפעוליים, הוצא מדגם דו-שלבי עבור האוכלוסייה ביישובים המונים פחות מ-8,000 בני 20 ומעלה. בשלב ראשון חולקו היישובים לשכבות גאוגרפיות והוצא מדגם של יישובים יחסית לגודלם. בשלב שני הוצא מדגם מקרי-שיטתי של נפשות ביישובים שנדגמו לאחר מיון הרשומות, לפי מאפייני 70 התכונות לתכנון. הסתברות הדגימה של השלב השני הותאמה כך שהסתברות הדגימה הסופית (כפל ההסתברויות בשני השלבים) תהיה קבועה. בשיטה זו התקבל גודל מדגם של כ-15 נפשות כנדרש. בסוף 2009 הוצא מדגם תוספות של עולים חדשים, אשר עלו החל מאפריל 2009 ולא נכללו במסגרת הדגימה העיקרית. בסך הכל כלל המדגם 9,539 נפשות במטרה לקבל כ-7,500 משיבים לסקר.

## שיטת הדגימה בסקר חברתי 2011

גודל המדגם תוכנן כך שמספר המשיבים הצפוי יהיה כ-7,500 בני 20 ומעלה. בסקר החברתי 2011 הוגדרו קבוצות תכנון שונות מהמחזוריים הקודמים של הסקר: אוחדו קבוצות התכנון של יהודים ילידי ישראל ויהודים שהם עולים ותיקים (עלו לפני 1990). בקבוצה המאוחדת לעיל, בגילים 20-64 (חמש קבוצות גיל), הייתה הבחנה בין חרדים ומי שאינם חרדים על פי מקור מינהלי שמבוסס בעיקר על מידע על מוסדות לימוד. לפיכך, המדגם תוכנן ל-66 קבוצות תכנון: 56 קבוצות בשילוב שלושת המשתנים: 4 קבוצות אוכלוסייה (ערבים ממזרח ירושלים, ערבים שלא ממזרח ירושלים, יהודים עולי 1990 ואילך, ויהודים ילידי ישראל או שעלו לישראל לפני 1990 ושאינם חרדים לפי המקור המנהלי),

2 קבוצות מין (גברים נשים) ו-7 קבוצות גיל (20-24, 25-34, 35-44, 45-54, 55-64, 65-74, +75).  
10 קבוצות נוספות נועדו עבור חרדים שאינם עולים חדשים, לפי מין ו-5 קבוצות גיל (20-64).

גודל מדגם המשיבים הצפוי לקבוצה תוכנן להיות יחסי לגודל הקבוצה (מספר הנפשות בקבוצת תכנון), באילוץ מספר משיבים צפוי - 7,500. על סמך ההנחות לשיעורי אי-השבה (נפל) ואי-שייכות לאוכלוסיית הסקר (אפס), בכל קבוצה התקבל הגודל הדרוש לדגימה. הסתברות הדגימה הממוצעת הייתה 1:542, הסתברות הדגימה המקסימלית הייתה 1:410 והמינימלית - 1:615.

הנפשות ביישובים שבהם היו צפויים להתקבל לפחות 15 מקרים במדגם (יישובים שבהם כ-8,100 בני 20 ומעלה), נדגמו במדגם שכבות חד-שלבי, כאשר כל שכבת דגימה היא קבוצת תכנון. בכל שכבה הוצא מדגם מקרי-שיטתי של נפשות, לאחר מיון הקובץ לפי מאפיינים גאוגרפיים. בסך הכל כ-83% מהמדגם נערך בשיטת דגימה חד-שלבית.

עקב אילוצים תפעוליים, הוצא מדגם דו-שלבי עבור האוכלוסייה ביישובים המונים פחות מ-8,100 בני 20 ומעלה. בשלב ראשון חולקו היישובים לשכבות גאוגרפיות והוצא מדגם של יישובים יחסית לגודלם. בשלב שני הוצא מדגם מקרי-שיטתי של נפשות ביישובים שנדגמו לאחר מיון הרשומות, לפי מאפייני 70 התכונות לתכנון. הסתברות הדגימה של השלב השני הותאמה כך שהסתברות הדגימה הסופית (כפל ההסתברויות בשני השלבים) תהיה קבועה. בשיטה זו התקבל גודל מדגם של כ-15 נפשות כנדרש. בסוף 2010 הוצא מדגם תוספות של עולים חדשים, אשר עלו החל מאפריל 2010 ולא נכללו במסגרת הדגימה העיקרית. בסך הכל כלל המדגם 9,533 נפשות במטרה לקבל כ-7,500 משיבים לסקר.

### **שיטת הדגימה בסקר חברתי 2012**

גודל המדגם תוכנן כך שמספר המשיבים הצפוי יהיה כ-7,500 בני 20 ומעלה. קבוצות התכנון בסקר 2012 היו זהות לאלו שנקבעו לקראת סקר 2011 (ושונות מהמחזוריים הקודמים של הסקר): אוחדו קבוצות התכנון של יהודים ילידי ישראל ויהודים שהם עולים ותיקים (עלו לפני 1990). בקבוצה המאוחדת לעיל, בגילים 20-64 (חמש קבוצות גיל), הייתה הבחנה בין חרדים ומי שאינם חרדים על פי מקור מינהלי שמבוסס בעיקר על מידע על מוסדות לימוד. מכאן, המדגם תוכנן ל-66 קבוצות תכנון. 56 קבוצות בשילוב שלושת המשתנים: 4 קבוצות אוכלוסייה (ערבים ממזרח ירושלים, ערבים שלא ממזרח ירושלים, יהודים עולי 1990 ואילך, ויהודים ילידי ישראל או שעלו לישראל לפני 1990 ושאינם חרדים לפי המקור המנהלי), 2 קבוצות מין (גברים ונשים) ו-7 קבוצות גיל (20-24, 25-34, 35-44, 45-54, 55-64, 65-74, +75). 10 קבוצות תכנון נוספות היו עבור חרדים שאינם עולים חדשים, לפי מין ו-5 קבוצות גיל (20-64). גודל מדגם המשיבים הצפוי לקבוצה תוכנן להיות יחסי לגודל הקבוצה (מספר הנפשות במסגרת בקבוצת תכנון), ותחת האילוץ של מספר משיבים צפוי - 7,500. על סמך ההנחות לשיעורי אי-השבה (נפל) ואי-שייכות לאוכלוסיית הסקר (אפס), בכל קבוצה התקבל הגודל הדרוש לדגימה. הסתברות הדגימה הממוצעת הייתה 1:551, הסתברות הדגימה המקסימלית הייתה 1:417 והמינימלית - 1:625.

הנפשות ביישובים שבהם היו צפויים להתקבל לפחות 15 מקרים במדגם (כלומר שבהם כ-8,300 נפשות בני 20 ומעלה), נדגמו במדגם שכבות חד-שלבי, כאשר כל שכבת דגימה היא קבוצת תכנון. בכל שכבה הוצא מדגם מקרי-שיטתי של נפשות לאחר מיון הקובץ לפי מאפיינים גאוגרפיים. בסך הכל, כ-82% מהמדגם הוצא בשיטה דגימה חד-שלבית.

כדי לענות על האילוצים התפעוליים, הוצא מדגם דו-שלבי עבור האוכלוסייה ביישובים המונים פחות מ-8,300 בני 20 ומעלה. בשלב ראשון חולקו היישובים לשכבות גאוגרפיות והוצא מדגם של יישובים יחסית לגודלם. בשלב שני הוצא מדגם מקרי-שיטתי של נפשות ביישובים שנדגמו לאחר מיון הרשומות לפי מאפייני 66 התכונות לתכנון. הסתברות הדגימה של השלב השני הותאמה כך שהסתברות הדגימה הסופית (כפל ההסתברויות בשני השלבים) תהיה קבועה. בשיטה זו התקבל גודל מדגם של כ-15 נפשות כנדרש. בסוף 2011 הוצא מדגם תוספות של עולים חדשים, אשר עלו החל מאפריל 2011 ולא נכללו במסגרת הדגימה העיקרית. בסך הכל כלל המדגם 9,516 נפשות במטרה לקבל כ-7,500 משיבים לסקר.

### שיטת הדגימה בסקר חברתי 2013

גודל המדגם תוכנן כך שמספר המשיבים הצפוי יהיה כ-7,500 בני 20 ומעלה. קבוצות התכנון בסקר 2013 היו זהות לאלו שנקבעו לקראת סקר 2012 (ושונות מהמחזורים הקודמים של הסקר): אוחדו קבוצות התכנון של יהודים ילידי הארץ ויהודים שהם עולים ותיקים (עלו לפני 1990). בקבוצה המאוחדת לעיל, בגילאים 20-64 (חמש קבוצות גיל), הייתה אבחנה בין חרדים ומי שאינם חרדים על פי מקור מנהלי שמבוסס בעיקר על מידע על מוסדות הלימודים. כמו כן הייתה אבחנה בין אקדמאים בוגרי אוניברסיטאות או סטודנטים ב-10 השנים האחרונות, לבין מי שאינו כזה בחלק מקבוצות הגיל\*מין\*קבוצת-אוכלוסייה. מכאן, המדגם תוכנן ב-86 קבוצות תכנון. 56 משילוב שלושת המשתנים: 4 קבוצות אוכלוסייה (ערבים ממזרח ירושלים, ערבים שלא במזרח ירושלים, יהודים עולי 1990 ואילך, ויהודים ילידי הארץ או שעלו לפני 1990 ואינם חרדים לפי המקור המנהלי), 2 קבוצות מין (גברים ונשים) ו-7 קבוצות גיל (20-24, 25-34, 35-44, 45-54, 55-64, 65-74, +75). 10 קבוצות תכנון נוספות היו עבור חרדים שאינם עולים חדשים, לפי מין ו-5 קבוצות גיל (20-64), ו-20 קבוצות תכנון של בוגרי אוניברסיטאות וסטודנטים ב-10 השנים האחרונות. גודל מדגם המשיבים הצפוי לקבוצה תוכנן להיות יחסית לגודל הקבוצה (מספר הנפשות במסגרת בקבוצת תכנון), ותחת האילוץ של מספר משיבים צפוי 7,500. על סמך ההנחות לשיעורי אי ההשבה (נקרא נפל) ואי שייכות לאוכלוסיית הסקר (נקרא אפס), בכל קבוצה התקבל הגודל הדרוש לדגימה. הסתברות הדגימה הממוצעת הייתה 1:551, הסתברות הדגימה המקסימאלית הייתה 1:423 והמינימאלית 1:627.

הנפשות ביישובים שבהם היו צפויים להתקבל לפחות 15 מקרים במדגם (כלומר שבהם כ-8,300 נפשות בני 20 ומעלה), נדגמו במדגם שכבות חד-שלבי, כאשר כל שכבת דגימה היא קבוצת תכנון. בכל שכבה הוצא מדגם מקרי-שיטתי של נפשות לאחר מיון הקובץ על ידי מאפיינים גאוגרפיים. בסך הכול, כ-83% מהמדגם הוצא בשיטה דגימה חד-שלבית.

כדי לענות על האילוצים התפעוליים, הוצא עבור האוכלוסייה ביישובים עם פחות מ-8,300 בני 20 ומעלה מדגם דו-שלבי. בשלב ראשון חולקו היישובים לשכבות גיאוגרפיות והוצא מדגם של יישובים יחסית לגודלם. בשלב שני הוצא מדגם מקרי-שיטתי של נפשות ביישובים שנדגמו לאחר מיון הרשומות לפי מאפייני 86 התכונות לתכנון. הסתברות הדגימה של השלב השני הותאמה כך שהסתברות הדגימה הסופית (כפל ההסתברויות בשני השלבים) תהיה קבועה. בשיטה זו התקבל גודל המדגם של כ-15 נפשות כנדרש. בסוף 2012 הוצא מדגם תוספות של עולים חדשים, אשר עלו החל מאפריל 2012 ולא נכללו במסגרת הדגימה העיקרית. בסך הכול כלל המדגם 9,604 נפשות במטרה לקבל כ-7,500 משיבים לסקר.

## שיטת הדגימה בסקר חברתי 2014

### שיטת האמידה

הנתונים הנאספים בסקר החברתי מבוססים על מדגם של נפשות מהאוכלוסייה. בכדי להפיק אומדנים על כלל האוכלוסייה ועל תת-הקבוצות שבה, יש לחשב לכל משיב בסקר מקדם ניפוח, המבטא את אומדן מספר הנפשות באוכלוסייה "המיוצגות" על ידו. שיטת האמידה קשורה לשיטת הדגימה ולבעיות שונות באיסוף הנתונים. אחת הבעיות השכיחות היא שלא כל מי שנדגם השיב לסקר ואלה שלא השיבו לסקר אינם בהכרח דומים בתכונותיהם לאלה שהשיבו (אי-השבה אינפורמטיבי) ולכן עלולה להיגרם הטיות באומדני הסקר. שיטת האמידה מיועדת לצמצם בעיקר את ההטיות העוללות לנבוע מאי-השבה אינפורמטיבית, מתת-כיסוי של האוכלוסייה ומשונות בגודלי המדגם לפי תכונות שלא נכללו בתכנון מדגם הסקר, כגון מאפייני עבודה והשכלה.

תהליך האמידה כולל שלושה שלבים: בשלב הראשון כל משיב קיבל מקדם ניפוח ראשוני שהוא היפוך הסתברות הדגימה, ומבטא את הדיפרנציאליות בתכנון המדגם; בשלב השני הוכפל מקדם הניפוח הראשוני במקדם תיקון לנפל חריג וזאת לרשומות באזורים גאוגרפיים שבהם התקבלו שיעורים חריגים במיוחד של אי-השבה; לשלב השלישי נבחרו מבין משתני הסקר משתני התאמה, אשר התפלגותם ידועה ממקורות מנהליים או נאמדת מסקר כוח אדם (סקר גדול אחר של הלמ"ס), ושנמצאו קשורים באופן החזק ביותר למשתני חקירה מרכזיים בשאלון הגרעין של הסקר. בחירת המשתנים נעשתה על סמך נתוני הסקר החברתי 2002. בשלב השלישי חושבו מקדמי הניפוח הסופיים בשיטת ה-Raking, שלפיה מתאימים את התפלגות המדגם המנופח במקדמי ניפוח של השלב הקודם לשלוש התפלגויות של אומדנים חיצוניים:

- אומדני משקי בית לפי קבוצת אוכלוסייה \* גודל משק בית מורחב – 37 קבוצות.
- אומדני נפשות בגיל +20 לפי קבוצת אוכלוסייה \* תכונות כוח עבודה \* השכלה - 12 קבוצות.
- אומדני נפשות בגיל +20 לפי קבוצת אוכלוסייה \* קבוצות גאוגרפיות \* מין \* קבוצות גיל - 191 קבוצות.

החל משנת 2012 המשתנה קבוצת האוכלוסייה ביישובים היהודים והמעורבים כולל הבחנה בין חרדים ממקורות מנהליים (לפי מוסדות חינוך) לבין יתר האוכלוסייה (עבור בני 20-64). בסקר החברתי 2012 מספר המשיבים ביישובי הבדווים בדרום היה 15, זאת עקב קשיים תפעוליים. לכן קביעת מקדמי הניפוח לסקר בשיטה שתוארה לעיל נעשתה בלעדיהם ואילו מקדמי הניפוח של הבדווים ביישובי הדרום התקבלו על ידי חלוקת האומדן הדמוגרפי ביישובים אלה במספר המשיבים וללא התניה בתכונותיהם.

בסקר 2013, בשלב השלישי חושבו מקדמי הניפוח הסופיים בשיטת ה-Raking לפיה מתאימים את התפלגות המדגם המנופח במקדמי הניפוח של השלב הקודם לארבע התפלגויות של אומדנים חיצוניים:

- אומדני משקי בית לפי קבוצת אוכלוסייה \* גודל מ"ב מורחב – 38 קבוצות.
  - אומדני נפשות בגיל +20 לפי קבוצת אוכלוסייה \* תכונות כוח עבודה \* השכלה - 11 קבוצות.
  - אומדני נפשות בגיל +20 לפי קבוצת אוכלוסייה \* קבוצות גיאוגרפיות \* מין \* קבוצות גיל - 230 קבוצות.
  - אומדני נפשות בגיל +20 לפי קבוצות אוכלוסייה \* קבוצות גיאוגרפיות \* הבחנה בין חרדים ממקורות מנהליים (על סמך מוסד לימודים אחרון), לבין יתר האוכלוסייה (בגיל 20-64), עבור האוכלוסייה ביישובים יהודיים ומעורבים – 26 קבוצות.
- בסקר החברתי 2013 לא נאספו נתונים ביישובי הבדווים בדרום, עקב קשיים תפעוליים. לכן קביעת מקדמי הניפוח לסקר בשיטה שתוארה לעיל נעשתה בלעדיהם.

## **איסוף נתונים**

עבודת השדה השנתית בסקר החברתי מתחילה מדי חודש ינואר ומסתיימת לרוב בחודש דצמבר. בשנת 2014 נמשכה עבודת השדה עד לחודש אפריל 2015. הסקר נערך על ידי מראיינים של הלמ"ס, בראיונות פנים-אל-פנים באמצעות מחשבים ניידים (CAPI), באמצעות תכנת Blaise, מפיתוחה של הלשכה הסטטיסטית ההולנדית. הראיונות נערכים בעברית, ברוסית ובערבית, ונמשכים כשעה. הנתונים נשמרים במחשב ומועברים מבית הסוקר למרכז האיסוף בירושלים באמצעות האינטרנט (בטלפון קווי או סלולארי).

## **סקר חברתי 2015**

### **שיטת הדגימה**

גודל המדגם תוכנן כך שמספר המשיבים הצפוי יהיה כ-7,500 בני 20 ומעלה. קבוצות התכנון בסקר 2015 היו זהות לאלו שנקבעו לקראת סקר 2013 וסקר 2014 (ושונות מהמחזוריים הקודמים של הסקר): אוחדו קבוצות התכנון של יהודים ילידי הארץ ויהודים שהם עולים ותיקים (עלו לפני 1990). בקבוצה המאוחדת לעיל, בגילאים 20-64 (חמש קבוצות גיל), הייתה אבחנה בין חרדים ומי שאינם חרדים על פי מקור מנהלי שמבוסס בעיקר על מידע על מוסדות הלימודים. כמו כן היתה אבחנה בין אקדמאים בוגרי אוניברסיטאות או



סטודנטים ב-10 השנים האחרונות, לבין מי שאינו כזה בחלק מקבוצות הגיל\*מין\*קבוצת-אוכלוסייה. מכאן, המדגם תוכן ב-86 קבוצות תכנון. 56 משילוב שלושת המשתנים: 4 קבוצות אוכלוסייה (ערבים ממזרח ירושלים, ערבים שלא במזרח ירושלים, יהודים עולי 1990 ואילך, ויהודים ילידי הארץ או שעלו לפני 1990 ואינם חרדים לפי המקור המנהלי), 2 קבוצות מין (גברים ונשים) ו-7 קבוצות גיל (20-24, 25-34, 35-44, 45-54, 55-64, 65-74, +75). 10 קבוצות תכנון נוספות היו עבור חרדים שאינם עולים חדשים, לפי מין ו-5 קבוצות גיל (-64, 20), ו-20 קבוצות תכנון של בוגרי אוניברסיטאות וסטודנטים ב-10 השנים האחרונות. גודל מדגם המשיבים הצפוי לקבוצה תוכן להיות יחסית לגודל הקבוצה (מספר הנפשות במסגרת בקבוצת תכנון), ותחת האילוץ של מספר משיבים צפוי 7,500. על סמך ההנחות לשיעורי אי ההשבה (נקרא נפל) ואי שייכות לאוכלוסיית הסקר (נקרא אפס), בכל קבוצה התקבל הגודל הדרוש לדגימה. הסתברות הדגימה הממוצעת הייתה 1:555, הסתברות הדגימה המקסימאלית הייתה 1:434 והמינימאלית 1:652.

הנפשות ביישובים שבהם היו צפויים להתקבל לפחות 15 מקרים במדגם (כלומר שבהם כ-8,300 נפשות בני 20 ומעלה), נדגמו במדגם שכבות חד-שלבי, כאשר כל שכבת דגימה היא קבוצת תכנון. בכל שכבה הוצא מדגם מקרי-שיטתי של נפשות לאחר מיון הקובץ על ידי מאפיינים גיאוגרפיים. בסך הכול, כ-83% מהמדגם הוצא בשיטה דגימה חד-שלבית.

כדי לענות על האילוצים התפעוליים, הוצא עבור האוכלוסייה ביישובים עם פחות מ-8,300 בני 20 ומעלה מדגם דו-שלבי. בשלב ראשון חולקו היישובים לשכבות גיאוגרפיות והוצא מדגם של יישובים יחסית לגודלם. בשלב שני הוצא מדגם מקרי-שיטתי של נפשות ביישובים שנדגמו לאחר מיון הרשומות לפי מאפייני 86 התכונות לתכנון. הסתברות הדגימה של השלב השני הותאמה כך שהסתברות הדגימה הסופית (כפל ההסתברויות בשני השלבים) תהיה קבועה. בשיטה זו התקבל גודל המדגם של כ-15 נפשות כנדרש. בסוף 2014 הוצא מדגם תוספות של עולים חדשים, אשר עלו החל מאפריל 2014 ולא נכללו במסגרת הדגימה העיקרית. בסך הכול כלל המדגם 9,603 נפשות במטרה לקבל כ-7,500 משיבים לסקר.

### **שיטת האמידה**

הנתונים הנאספים בסקר החברתי מבוססים על מדגם של נפשות מהאוכלוסייה. בכדי להפיק אומדנים עם תכונות נחקרות מהסקר לכלל האוכלוסייה ולתת-קבוצות באוכלוסייה, יש לחשב לכל משיב בסקר מקדם ניפוח, המבטא את אומדן מספר הנפשות באוכלוסייה "המיוצגות" על ידו. שיטת האמידה קשורה לשיטת הדגימה ולבעיות שונות באיסוף הנתונים. אחת הבעיות השכיחות היא שלא כל מי שנדגם השיב לסקר ויחידות אשר לא השיבו לסקר אינן בהכרח דומות בתכונותיהם ליחידות המשיבות (אי השבה אינפורמטיבי) ולכן עלולות לגרום להטיה באומדני הסקר. שיטת האמידה מיועדת לצמצם בעיקר את ההטיות העלולות לנבוע מאי

השבה אינפורמטיבית, מתת-כיסוי של האוכלוסייה ומשונות בגודלי המדגם לפי תכונות אשר לא השתתפו בתכנון מדגם הסקר, כמו למשל, תכונות עבודה והשכלה.

תהליך האמידה כולל שלושה שלבים: בשלב ראשון כל משיב קיבל מקדם ניפוח ראשוני השווה למכפלת היפוך הסתברות הדגימה בגורם תיקון לאי ההשבה. מקדם הניפוח הראשוני מבטא את הדיפרנציאליות בתכנון המדגם ואת דיפרנציאליות ההשבה בחתך גיאוגרפי צר; בשלב השני נבחרו משתני התאמה, מבין משתני הסקר אשר התפלגותם ידועה ממקורות מנהליים או נאמדת מסקר כוח-אדם חודשי (סקר גדול אחר של הלמ"ס), להם נמצא הקשר החזק ביותר עם אוסף משתני חקירה מרכזיים בשאלון הגרעין של הסקר. בשלב השלישי חושבו מקדמי הניפוח הסופיים בשיטת ה-Raking לפיה מתאימים את התפלגות המדגם המנופח במקדמי הניפוח של השלב הקודם לחמש התפלגויות של אומדנים חיצוניים:

1. אומדני משקי בית לפי קבוצת אוכלוסייה \* גודל מ"ב מורחב – 38 קבוצות.
2. אומדני נפשות בגיל +20 לפי קבוצת אוכלוסייה \* מין \* מצב משפחתי – 12 קבוצות.
3. אומדני נפשות בגיל +20 לפי קבוצת אוכלוסיה \* תכונות כוח עבודה \* השכלה – 11 קבוצות.
4. אומדני נפשות בגיל +20 לפי קבוצת אוכלוסייה \* קבוצת גיאוגרפיות \* הבחנה בין חרדים ממקורות מנהליים (על סמך מוסד לימודים אחרון), לבין יתר האוכלוסייה (בגיל 20-64), עבור האוכלוסייה ביישובים יהודיים ומעורבים – 26 קבוצות.
5. אומדני נפשות בגיל +20 לפי קבוצת אוכלוסייה \* קבוצת גיאוגרפיות \* מין \* קבוצת גיל - 233 קבוצות.

### **חשיבות האמידה ושימוש במקדמי הניפוח**

הנתונים הנאספים בסקר החברתי מבוססים על מדגם של נפשות מהאוכלוסייה. בכדי להפיק אומדנים עם תכונות נחקרות מהסקר לכלל האוכלוסייה ולתת-קבוצות באוכלוסייה, חושב לכל משיב בסקר מקדם ניפוח, המבטא את אומדן מספר הנפשות באוכלוסייה "המיוצגות" על ידו. השדה "מקדם ניפוח" שבקובץ (המשתנה חח) משמש לקבלת אומדן לאוכלוסיית הסקר - בני 20 ומעלה.

### **מהימנות האומדנים**

מקובץ נתוני הסקר ניתן להפיק אומדני סך הכל ואומדני אחוזים עם שימוש במשקל (מקדם הניפוח). בפרסומים של הלמ"ס, כולל במחולל הלוחות של הסקר החברתי, בצד האומדנים ניתן לחשב את טעויות הדגימה וטעויות היחסיות (טעות הדגימה חלקי האומדן). טעות הדגימה היחסית מאפשרת לנו להיזהר מאומדנים עם רמות מהימנות נמוכות ונמוכות מאד. משתמש שאינו מחשב טעויות דגימה יכול לקבל קירוב מסוים על ידי גודל המדגם המשתתף בחישוב האומדן.

מניסיון הלמ"ס עבור אומדני סך הכל: אומדן המבוסס על לכל היותר 11 מקרים – הוא בעל רמת מהימנות נמוכה מאד ואין להשתמש בו. אומדן המבוסס על 12 עד 44 מקרים – יש להתייחס אליו כבעל רמת מהימנות נמוכה ואילו אומדן המבוסס על 45 מקרים ויותר הוא מקביל לאומדן אשר הלשכה נוהגת לפרסם ללא הערת אזהרה.

לגבי אומדני אחוזים המלצתנו היא: אומדן אחוז אשר המכנה מבוסס על 25 מקרים לכל היותר - אם המונה הוא 1-5 מקרים נגדיר אותו כבעל רמת מהימנות נמוכה מאד ואין להשתמש בו, אם במונה 6 עד 10 מקרים יש להתייחס אליו כבעל רמת מהימנות נמוכה ואם מעל 10 מקרים ללא אזהרה. לעומת זאת, במקרים אשר המכנה מבוסס על יותר מ-25 מקרים, אם המונה הוא עד 8 מקרים נגדיר אותו כבעל רמת מהימנות נמוכה מאד ואין להשתמש בו, אם במונה 9 עד 12 מקרים יש להתייחס אליו כבעל רמת מהימנות נמוכה ואם מעל 12 מקרים ללא אזהרה.