

סקר חברתי 2010 PUF מתודולוגית הסקר החברתי

אוכלוסיית הסקר

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

שיטת הדגימה

סקר החברתי 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) וחלק מתושבי המוסדות ותושבי יישובים מסוימים. כדי לייעל את אומדני הסקר ותהליך איסוף הנתונים, זוהו ואותרו חלק מהרשומות העודפות לפני הוצאת המדגם, ונכנסו מקובץ מרשם התושבים ששימש כמסגרת לדגימה. הדגימה התבצעה בנפרד עבור יישובים גדולים ויישובים קטנים. יישובים גדולים הוגדרו כיישובים עם 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, 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 גיאוגרפיים 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 בדגימה לסקר החברתי 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 המינימאלי ביישוב גדל מ-10 נפשות ל-15 נפשות. כתוצאה מכך יישובים גדולים בהם מוצא מדגם שכבות חד-שלבי הם אלו עם 7,100 או יותר בני 20 ומעלה (כ-83% מאוכלוסיית הסקר מתגוררים ביישובים אלה).

שיטת הדגימה לסקר החברתי 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 במזרח ירושלים, ערבים שלא במזרח 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 הכול 70 קבוצות תכנון. גודל מדגם המשיבים הצפוי לקבוצה תוכנן להיות יחסית לגודל הקבוצה (מספר הנפשות במסגרת בקבוצת תכנון), ותחת האילוץ של מספר משיבים צפוי - 7,200. על סמך ההנחות לשיעורי אי-ההשבה (נקרא נפל) ואי-שייכות לאוכלוסיית הסקר (נקרא 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, על סמך המידע ממרשם התושבים, לכדי 34% מכלל המדגם, לעומת 29% אילו לא נערך תגבור כלל. הסתברות הדגימה הסופית לנפש משתנה בין קבוצות התכנון ולעיתים גם בתוך קבוצות התכנון והיא מבטאת הערכות מוקדמות לאחוזי אי-ההשבה בקבוצות השונות וההחלטות על תגבור ההורים. הסתברות הדגימה הממוצעת הייתה 1:485, הסתברות הדגימה המקסימאלית הייתה 1:305, והמינימאלית 1:920.

נפשות ביישובים שלהם 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 על ידי מאפיינים גיאוגרפיים. בחלק משכבות הדגימה נקבעה הסתברות הדגימה לנפש בהתאם למידע מהמרשם אם הנפש היא 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 הכול כ-83% מהמדגם הוצא בשיטת דגימה חד-שלבית.

כדי לענות על האילוצים התפעוליים, הוצא עבור האוכלוסייה ביישובים עם פחות מ-7,400 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 70 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.

בתחילת 2004 הוצא מדגם תוספות של עולים חדשים, אשר עלו לארץ החל מאפריל 2003 ולא נכללו במסגרת הדגימה העיקרית.

בסך הכול כלל המדגם 9,528 נפשות במטרה לקבל כ-7,200 משיבים לסקר.

שיטת הדגימה לסקר החברתי 2005

המדגם הוצא בשיטה דומה לזו ששימשה לסקר חברתי 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 לסקר החברתי 2006

המדגם התבסס על 70 קבוצות תכנון, הזהות 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, דבר שאפשר לאפיין מתוך המרשם את בני משפחתם הקרובה כ"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 הכול נדגמו לסקר 9,499 בני 20 ומעלה.

שיטת הדגימה לסקר החברתי 2007

המדגם הוצא בשיטה דומה 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 לסקר 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 הכול נדגמו לסקר 9,389 בני 20 ומעלה.

שיטת הדגימה לסקר החברתי 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 הכול נדגמו לסקר 9,340 בני 20 ומעלה. בדיעבד, התברר כי בשנת 2009 נדגמו מספר גדול יחסית של יישובים לא עירוניים יהודיים עם אוריינטציה דתית, אשר הביאו לאומדן גבוה יחסית של פרטים שהגדירו עצמם כדתיים בשאלון הסקר, זאת בהשוואה לשנת 2008.

שיטת הדגימה לסקר החברתי 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 במזרח ירושלים, ערבים שלא במזרח ירושלים, עולי 1990 והלאה, 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) – סך הכול 70 קבוצות תכנון. 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 1:531, הסתברות הדגימה המקסימאלית הייתה 1:380 והמינימאלית 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 גיאוגרפיים. בסך הכול, כ-84% מהמדגם הוצא בשיטה דגימה חד-שלבית.

כדי לענות על האילוצים התפעוליים, הוצא עבור האוכלוסייה ביישובים עם פחות מ-8,000 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 70 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 2009 הוצא מדגם תוספות של עולים חדשים, אשר עלו החל מאפריל 2009 ולא נכללו במסגרת הדגימה העיקרית. בסך הכול כלל המדגם 9,539 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 שהוא היפוך הסתברות הדגימה, המבטא את הדיפרנציאליות בתכנון המדגם; בשלב שני הוכפל מקדם הניפוח הראשוני במקדם תיקון לנפל חריג וזאת לרשומות באזורים גיאוגרפיים בהם התקבלו שיעורי אי השבה חריגים במיוחד; לביצוע השלב השלישי נבחרו משתני התאמה, מבין משתני הסקר אשר התפלגותם ידועה ממקורות מנהליים או נאמדת מסקר כוח-אדם (סקר גדול אחר של הלמ"ס), להם נמצא הקשר החזק ביותר עם אוסף משתני חקירה מרכזיים בשאלון הגרעין של הסקר. בחירת המשתנים נעשתה על סמך נתוני הסקר החברתי 2002. בשלב השלישי חושבו מקדמי הניפוח הסופיים בשיטת ה-Raking לפיה מתאימים את התפלגות המדגם המנופח במקדמי הניפוח של השלב הקודם לשלוש 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 - 15 קבוצות.
- אומדני נפשות בגיל +20 לפי קבוצת אוכלוסייה * קבוצות גיאוגרפיות * מין * קבוצות גיל - 198 קבוצות.

איסוף נתונים

עבודת השדה השנתית בסקר החברתי מתחילה מדי חודש ינואר ומסתיימת בחודש דצמבר. הסקר נערך על ידי מראיינים של הלמ"ס, בראיונות פנים-אל-פנים באמצעות מחשבים ניידים (CAPI), באמצעות תוכנת 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).