

נניח ואנחנו רוצים לשאול שאלה מסוימת- מי ינצח בבחירות? ברור שלא נשאל את כל האוכלוסייה שהולכת לקלפי להצביע (למרות שחוסכים את טעות הדגימה אני אתקשה לבצע את זה בגלל: זמן, עלות ולא מדויק בהכרח- תפעול גדול מאוד שמצריך פיקוח רב שלא תמיד מתאפשר). לכן, משתמשים בדגימה. הסיבה העיקרית שבגללה אני משתמשת בדגימה היא כדי לחסוך בזמן.
מתי לא עושים דגימה? במפקד אוכלוסין.
מפקד אוכלוסין הוא סקר שבו משתתפת כל האוכלוסייה. בכל מדינה מתוקנת מתקיים כל עשר שנים מפקד אוכלוסין בשביל לספור את האוכלוסייה באותה המדינה. מחפשים את הבדיקה הכי מדויקת שיש. זהו סקר שמגובה בחוק- חייבים לענות על הסקר ולא לשקר. מי שמשקר או לא עונה על הסקר- יועמד לדין.
מתי כדאי לעשות מפקד?

- כאשר האוכלוסייה היא קטנה באופן מיוחד- חברה שמתעסקת בצידוד משרדי שרוצה לבחון את כל החנויות שאליהן היא משוקקת ויש לה סה"כ 300 חנויות כאלו. במקרה כזה שווה לעשות 100% דגימה.

- כאשר יש צורך לקבל מידע מדויק ככל הניתן ולהתגבר על טעויות דגימה.

- כאשר רוצים לקבל מידע על תאים סטטיסטיים קטנים.

מתי כדאי לדגום?

- כאשר האוכלוסייה היא גדולה

- כאשר יש בעיית זמן

- כאשר עלות איסוף המידע מכולם היא גבוהה

סיפור על בית אבות - הסיפור נמצא בשקף שבמצגת שעל האתר. הסיפור מתאר מפקד- כל מי שמגיע. אך למעשה האוכלוסייה מוטעית- זו אוכלוסייה שכנראה שמעה רק את הדברים הטובים על המקום, או שהקבוצה הזו סובלת מדיסוננס קוגניטיבי (נובע מהצורך של אנשים לא להרגיש מטומטמים. אנשים רוצים לשדר שהחלטה שלהם היא ההחלטה הכי טובה שהם יכלו לקבל)- האנשים בבית האבות צריכים לשכנע את קרוביהם שהם נמצאים במקום הכי טוב שיש.

סיפור על משרד שליחויות - הסיפור נמצא בשקף שבמצגת שעל האתר. זהו מצב קלאסי של גוף שחושב שע"י זה שהוא חושב שיש לו 10,000 שאלונים זה הופך למייצג. האנשים שענו הם אנשים מאוד ספציפיים- או שהם מאוד משועממים, או מאוד אוהבים את הבנק או מאוד מפחדים מהבנק. צריך לשים לב ולא לתת למספרים הגדולים לבלבל אותנו.

סיפור על משרד קייטרינג - הסיפור נמצא בשקף שבמצגת שעל האתר. אם השאלון הוא טלפוני ובערב- מקבלים הטיה של אופי ההתנהגות של האחיות.

הגדרות:

- אוכלוסייה- כל היחידות (סוגי עסקים, אנשים), שיש להן כמה תכונות משותפות שקשורות לבעיית מחקר השוק. אני מגדירה

- דגימה- בחירה של חלק מהיחידות מתוך האוכלוסייה

- יחידות הדגימה- units sampling - היחידות אותן אנחנו דוגמים (אנשים, משקי בית, חנויות, בתי ספר וכיו"ב). המטרה היא למנוע כפל דגימה או לפספס נדגמים שהחלטנו לדגום. לתת הגדרה מדויקת ל: מהו "משק בית"? אבא, אמא או ילד. אחת השיטות היא לבדוק מי מבצע את הפעולה שמעניינת אותנו.

שתי דרכים להחליט מהן יחידות הדגימה:
הדרך המדויקת- מבקשים שמי שיענה לסקר הוא מי שיום ההולדת שלו הכי קרוב לתאריך מסויים.
בנוסף, בסקרים טלפוניים עושים מכסות (50% גברים ו- 50% נשים)- אם סקרנו כבר 50% גברים, מעכשיו
נבקש לדבר רק עם הנשים.

- מסגרת הדגימה- frame sampling - רשימה שבה נמצאות כל יחידות הדגימה שממנה ניתן לדגום
(ספר טלפונים, רשימת ארגונים, רשימת סטודנטים, רשימת רחובות, רשימת יישובים וכו').

תהליך הדגימה:

- הגדרת האוכלוסייה- שלב קריטי בכל עבודה מחקרית. את מי אני בדיוק בודקת. זוהי הגדרה מאוד
חשובה .

- איך קובעים את אוכלוסיית המחקר?
א. להיצמד ליעדי המחקר (שאלות וגבולות המחקר)
ב. לשקול אלטרנטיבות- אם אני רוצה לפתוח חנות צעצועים- עלי לשקול את מי אני
דוגמת: רק הורים? רק ילדים? גם וגם?
ג. להתבסס על מידע אודות השוק (מחקרים קודמים, מחקרי גישוש)
ד. להגדיר באופן ברור מה לא נכלל באוכלוסיית המדגם

- הגדרת מסגרת הדגימה- המסגרת היחידה שיש כיום בארץ זה הרשימות של בזק ושל חברת [ביטוח רכב](#) וליסנג. עם הרשימות הללו יש שתי בעיות: א) ישנם אנשים שלא כתובים ברשימות של בזק ב)לאנשים יש יותר מקו טלפון אחד בבית.

* איך מתקנים מצב של אנשים שאינם רשומים ברשימות של בזק? בכל חברת טלפונים יש רצף של מספרים פעילים. אז מה שעושים זה מתקשרים באופן אקראי למספר מתוך הרצף.
* איך מתקנים מצב של הטיה לגבי אנשים שיש להם יותר מקו טלפון אחד בבית? שואלים אנשים כמה קווי טלפון פעילים יש להם בבית. ובסוף המדגם משקללים את הנתון הזה.

- בחירת סוג הדגימה- מדברים בגדול על שני סוגי דגימות:

א) דגימה הסתברותית- מוגדרת כדגימה שבה לכל יחידה יש הסתברות ידועה מראש להידגם. דגימה טובה יותר ויותר מייצגת.

דגימה הסתברותית:

(1) דגימה אקראית פשוטה - זוהי הדוגמה הפשוטה ביותר. לכל אחת מהיחידות יש הסתברות שווה להיבחר למדגם. בדר"כ זוהי הדגימה הכי פופולארית, מבוצעת ע"י מחשב או רשימות אקראיות שממספרים.
(2) דגימה שיטתית - אנחנו מגדירים מראש איזשהו אינטרוואל של דגימה (K). למשל $20=K$ אז אני אבחר בפעם הראשונה את היחידה ה- 20 ובפעם השנייה את היחידה ה- 40 וכו'. בדגימות של דלת לדלת- אומרים לסוקר לבחור בכל פעם את הדלת ה-6 (לדוגמא). משתמשים כדי ליצור אקראיות שלא ע"י מחשב.

(3) דגימת שכבות - כאן אנו מניחים / יודעים שאוכלוסיית המחקר בנויה משכבות יחסית הומוגניות לגבי מה שאנחנו עומדים לבדוק. נניח שאין קשר בין הצבעה בבחירות לבין מצב סוציו-אקונומי. אם אני רוצה לבדוק את דפוסי ההצבעה לראש העיר ת"א - אני יכול לדגום את האוכלוסייה בין צפון לדרום העיר. השאלה היא האם זה בכלל רלוונטי לתשובות שאמורות להתקבל. היות ולא נוכל להניח הומוגניות של

השכונות הצפוניות והדרומיות באופן ברור לא נוכל לפעול לפי דגימת השכבות. אם אני רוצה לבדוק אפיוני צריכה של תושבי העיר - כאן אני יכול להניח שיש שכבות באוכלוסייה וזה מבוסס הרבה מאוד על בסיס מקום מגורים. כאן אני משתמש בדגימת שכבות. יש לי כאן למעשה יתרון סטטיסטי - אנו מקבלים יתרון שמאפשר לנו להוריד את טעות הדגימה או להקטין את המדגם בלי לפגוע בטעות הדגימה (אם כולם אומרים אותו דבר - אני יכול לשאול רק את אחד הנדגמים וזה חוסך המון עלויות).

איך מזהים הומוגניות ? (ההומוגניות חשובה רק לעניין הנושא שאנחנו רוצים לבדוק)

1. לפי הנחות לוגיות (אפשר להניח שכל האוכלוסייה בדרום ת"א תהיה פחות או יותר זהה)
 2. שימוש בנתוני עבר
 3. שימוש במחקר גישוש
- נניח שאני מזהה 3 שכבות באוכלוסייה

(ב) דגימה לא הסתברותית- לא יודעים מראש מה ההסתברות של כל יחידה להידגם. דגימה שבהגדרתה היא מוטה.

דגימות לא הסתברותיות - נעשה שימוש בדגימה כזו בעיקר במצב שאין לנו ברירה (אין לנו רשימת דגימה):

(1) דגימה שיפוטית- אפשר להתבסס על נתונים לא לגמרי מדויקים. לדוגמא- כאשר מריצים פיילוט לשאלון. עושים את בהנחה וההתנהגות האנושית היא אוניברסאלית. בדגימה כזו אנחנו קובעים מראש (בצורה שיפוטית) קריטריון למדגם. עורכים מעין מבחן "שיפוט" כדי לקבוע מי יידגם: מידות קטנות או מידול גדולות, אנשים שמתלבשים בצורה אופנתית.

(2) כדור שלג - דגימה שמשתמשים בה המון. זו היא דגימה שבה מבקשים מאנשים שנמצאים בקבוצה מאוד ספציפית שיתנו שמות של אנשים נוספים שעונים על אותו קריטריון. דגימה שהיא מאוד אפקטיבית באוכלוסיות מיוחדות.

(3) דגימת נוחות - דגימה שהיא הכי נוחה והכי זולה להגיע לאנשים (ככה נעשה את השאלונים שלנו). לדוגמא: שאלונים בבתי מלון. הכוונה היא לדגום איפה שאנשים נמצאים - לא צריך להתאמץ כדי להגיע לאנשים, דוגמא נוספת- דגימת קניונים (הבעיה עם דגימת קניונים שיש הטיה- אני דוגמת יותר את מי שמגיע יותר. בעיה נוספת היא שהסוקרים פונים למי שהולך לאט יותר, מחייך יותר).

(4) דגימת מכסות- מחליטים שאנחנו רוצים X אנשים מכל קבוצת אוכלוסייה.