

ЕТАПИ НА СЪЧЕТА- ВАНЕТО НА ДАННИ

Съчетаването на данни преминава през няколко етапа, всеки от които има решаващо значение за качеството на получените крайни резултати. Про-

цесът започва с установяване на потенциалните източници, които подлежат на съчетаване. Най-общо те могат да се класифицират в три основни групи: административни източници (например регистри), извадкови изследвания и изчерпателни статистически изследвания.

Съществен въпрос, от отговора на който се определя дали изобщо е възможно съчетаването на данни, е установяването на статистическата съвместимост на източниците. Преди същинското съчетаване на данните трябва да се провери и ако не е налице, да се направи опит за постигане на хармонизиране на единиците, признаците, времевите и пространствените обхвати, използваните класификации и номенклатури, приложените методи за приписване на липсващите стойности и др. Хармонизирането на понятията и използването на стандартни класификации в различните източници има жизненоважно значение за успешното съчетаване на данни. У нас например се наблюдават различия в дефинициите и начина за установяване на безработицата при отделните лица в статистиката на домакинските бюджети и Наблюдението на работната сила. Това налага използването на известни допускания при съчетаването на тези източници, както и при съчетаването на данни от преброяването на населението и Наблюдението на работната сила поради същата причина.

Важен етап е логическият и физическият контрол на данните в източниците, които ще бъдат съчетавани. Това е доста трудоемък процес, качеството на който зависи от наличните метаданни за статистическите изследвания. Редактирането на данни е пряко свързано с проблемите на липсващите данни и опитът в това направление играе ключова роля.

Методологичните подробности за начините на същинско свързване на данните излизат извън обсега на настоящия доклад. Тук по-скоро е важно да се обърне внимание на следващия етап - оценка на качеството на постигнатото съчетаване. Три са основните елементи, от които зависи качеството на съчетаването на данни: 1) качеството на данните в източниците, които ще се съчетават; 2) свойствата на метода за съчетаване и 3) методите за изчисляването на производни статистически показатели от съчетаните данни. В литературата няма единна методология за измерване на качеството на съчетаването на данни. Продължават емпиричните опити за установяване на свойствата на оценките, получени въз основа на съчетани данни като неизместеност и ефективност. По-голяма част от проверките за качество се състои в провеждането на Монте Карло симулации за установяване