

3 Výběr a metodika dopočtu

3.1 Základní a výběrový soubor

Základní soubor pro strukturální šetření vychází z Registru farem (47 416 aktivních subjektů) a byl vytvořen na základě prahových hodnot (viz Definice zemědělského subjektu).

Z celkem 47 416 aktivních zemědělských subjektů vedených v Registru farem bylo do základního souboru vygenerováno 26 735 jednotek. V součtu obhospodařují 99,13 % celkové zemědělské půdy a chovají 98,70 % dobytčích jednotek (DJ).

Výběrový soubor (VS) byl stanoven jako stratifikovaný náhodný výběr ze základního souboru na základě příslušnosti k definované dopočtové skupině (stratu). Každé dopočtové skupině bylo přiřazeno určité procento výběru z celkového počtu jednotek v dané skupině (stratu) v rámci základního souboru. Jednotky s nižším procentem výběru než 100 % byly ze základního souboru generovány náhodným výběrem bez opakování. K tomuto účelu byl využit generátor náhodných čísel v programu SAS Enterprise Guide 4.2.

3.2 Stratifikace a dopočet

Zpravodajské jednotky byly zařazeny do strat, která byla kombinací tří proměnných:

- velikostní kategorie,
- příslušnost ke kraji sídlem farmy nebo bydlištěm fyzické osoby,
- příslušnost k zemědělské výrobní podoblasti.

Kód strata je tvořen šestiznakovým řetězcem tvaru XXYYZZ, kde:

- XX nabývá hodnoty 11 až 15 a 99 a rozlišuje velikost subjektu na základě výměry zemědělské půdy a/nebo dobytčích jednotek:
 - 11 (% výběru: 8 %): 0,00–9,99 ha zem. půdy nebo 0–1,99 DJ
 - 12 (% výběru: 11 %): 10,00–24,99 ha zem. půdy nebo 2–4,99 DJ
 - 13 (% výběru: 17 %): 25,00–49,99 ha zem. půdy nebo 5–9,99 DJ
 - 14 (% výběru: 21 %): 50,00–119,99 ha zem. půdy nebo 10–19,99 DJ
 - 15 (% výběru: 27 %): 120,00–199,99 ha zem. půdy nebo 20–49,99 DJ
 - 99 (% výběru: 100 %): - 200,00 a více ha zem. půdy
50 a více DJ
40 a více ks drůbeže
20 a více ks ovcí
5 a více ks koz
1,50 a více ha vinic
2,00 a více ha sadů
15,00 a více ha chmelnic
1,10 a více ha zeleniny
0,30 a více ha květin
0,30 a více ha jahod

Pro zařazení do jednotlivých velikostních kategorií mělo dosažení minimálního počtu kusů v kategorii 99 přednost před vyjádřením prostřednictvím zemědělské půdy či dobytčích jednotek.

- YY je zkrácený kód kraje a nabývá hodnot: 11, 21, 31, 32, 41, 42, 51, 52, 53, 61, 62, 71, 72 a 81.
- ZZ značí příslušnost k výrobní zemědělské podoblasti:
 - oblast kukuřičná s podoblastmi 11, 12, 13,
 - oblast řepařská s podoblastmi 21, 22, 23,

- o oblast bramborářská s podoblastmi 31, 32,
- o oblast bramborářsko-ovesná s podoblastí 33,
- o oblast horská s podoblastmi 41, 42.

Vzhledem k celkové malé výměře podoblastí 11, 12 a 13 byla jednotkám hospodařícím v podoblasti 12 a 13 přidělena pro výběr i dopočty podoblast 11. Obdobně byla pak nahrazena podoblast 42 podoblastí 41. Přípustné jsou tedy kódy podoblastí 11, 21, 22, 23, 31, 32, 33 a 41.

Počet takto zkonstruovaných dopočtových strat dosáhl hodnoty 521 kombinací, přestože teoreticky bylo možné dosáhnout celkem 672 (6 x 14 x 8) kombinací. V tomto členění byl stratifikační kód využit pouze pro generování VS (zachování proporcionality mezi kraji) a pro finální agregaci podle krajů. Při finálním dopočtu na ZS byl využit kód ve tvaru XXZZ, tedy maximálně 48 dopočtových strat.

Přepočtový koeficient každého strata K_p je dán vztahem:

$$K_p = N_z / n_v, \text{ kde}$$

K_p je přepočtový koeficient

N_z je počet jednotek základního souboru vytvořeného z registru farem za dané stratum

n_v je počet jednotek výběrového souboru za dané stratum

Výstupem dopočtu je agregovaný soubor vzniklý z datového souboru, jehož věty tvoří agregáty za jednotlivé dopočtové skupiny.

Poznámka: Při výpočtu směrodatné chyby bylo nalezeno 30 jednotek s extrémními hodnotami, jejichž extrapolací koeficient byl změněn (rovná se 1). Tyto jednotky byly přeřazeny do 100% strata.

3.3 Výstupní tabulky

Zpracované tabulky v publikaci sumarizují dopočtené údaje na základní soubor (zemědělství ČR) za zemědělské subjekty, které obhospodařují alespoň 98 % zemědělské půdy nebo chovají alespoň 98 % z celkového počtu dobytčích jednotek.

Publikované výsledky jsou dopočteny z nezaokrouhlených hodnot (data byla dopočítávána s přesností na 6 desetinných míst). Vyčerpávající součet v tabulkách se nemusí vždy přesně rovnat součtu svých složek, tak jak jsou v zaokrouhlené podobě uvedené v tabulkách.



Výpočty v tabulkách jsou prováděny z nezaokrouhlených údajů (vč. součtů).

POUŽITÉ ZNAČKY V TABULKÁCH PUBLIKACE

- ležatá čárka na místě čísla značí, že se jev nevyskytoval
- . tečka na místě čísla značí, že údaj není k dispozici nebo je nespolehlivý
- x ležatý křížek na místě čísla značí, že zápis není možný z logických důvodů
- 0 nula se používá pro označení číselných údajů menších než polovina zvolené měřicí jednotky
- i. d. individuální data