

Metodologia badania

Bilans
Kapitału
Ludzkiego

2016-2023



Metodologia badania
Bilans Kapitału Ludzkiego 2016-2023

Autorka Raportu:

Patrycja Antosz

© Copyright by Polska Agencja Rozwoju Przedsiębiorczości

© Copyright by Uniwersytet Jagielloński

ISBN 978-83-7633-405-9

Publikacja dostępna również na stronie internetowej
www.parp.gov.pl

Wersja elektroniczna

Warszawa 2018

Opracowanie graficzno-techniczne:
www.ccpog.com.pl



Unia Europejska
Europejski Fundusz Społeczny



Spis treści

Informacje o projekcie	4
Organizacja raportu.	6
Badania pracodawców	7
Podstawowe założenia	7
Populacja	8
Operat.	9
Dobór próby	9
Badania główne.	10
Wagi	11
Badania osób dorosłych	13
Podstawowe założenia	13
Populacja	14
Operat.	14
Dobór próby	14
Wagi	17
Badania instytucji i firm świadczących usługi rozwojowe	18
Podstawowe założenia	18
Populacja	19
Operat.	20
Dobór próby	22
Wagi	23

Informacje o projekcie

Bilans Kapitału Ludzkiego realizowany w latach 2016–2023 jest kontynuacją pierwszej edycji badań realizowanych w latach 2010–2015. Wówczas w różnych ośrodkach administracji centralnej oraz lokalnej pojawiło się zapotrzebowanie na szeroki zakres informacji związanych z zasobami kapitału ludzkiego dostępnego na polskim rynku pracy. W ramach pięciu edycji badań zrealizowanych w tamtym okresie zebrano szereg cennych informacji dotyczących m.in. zapotrzebowania pracodawców na określonych pracowników, sposobów rekrutacji, wymagań kompetencyjnych stawianym kandydatom, rozwoju pracowników, poziomu kompetencji posiadanych przez polskich pracowników, jak również bezrobotnych, osób nieaktywnych, uczniów i studentów, działań rozwojowych podejmowanych przez Polaków, oferowanych przez firmy i instytucje szkoleniowe działań rozwojowych. Informacje te były wielokrotnie wykorzystywane przez różne instytucje międzynarodowe (Komisja Europejska, CEDEFOP), centralne (np. Ministerstwo Rozwoju, Ministerstwo Nauki i Szkolnictwa Wyższego), lokalne (wojewódzkie urzędy pracy i urzędy marszałkowskie) oraz uczelnie. Opracowane wskaźniki monitorujące rynek pracy stały się swojego rodzaju standardem. To wszystko spowodowało konieczność kontynuowania badań w zmodyfikowanej formie, umożliwiającej jeszcze lepszy pomiar zjawisk związanych z rynkiem pracy i zmianami jakim on podlega. W związku z tym w edycjach badań zaplanowanych na lata 2016–2023 przewidziano pogłębienie niektórych wskaźników (np. związanych z rozwojem kompetencji pracowników) oraz panelowy charakter badań pracodawców oraz osób dorosłych, umożliwiający lepsze uchwycenie zmian tych wskaźników. Podobnie jak poprzednio projekt realizowany jest przez Polską Agencję Rozwoju Przedsiębiorczości w partnerstwie z Centrum Ewaluacji i Analiz Polityk Publicznych Uniwersytetu Jagiellońskiego.

Głównym celem projektu jest zdiagnozowanie oraz monitorowanie zasobów kapitału ludzkiego istniejących na polskim rynku pracy oraz czynników odpowiadających za ich rozwój. Cel ten może być osiągnięty poprzez realizację celów szczegółowych polegających na realizacji badań: pracodawców, osób dorosłych oraz instytucji i firm świadczących usługi rozwojowe.

Badania pracodawców:

- określenie zapotrzebowania na pracowników,
- określenie zapotrzebowania na kompetencje,

- określenie strategii zarządzania kapitałem ludzkim (świadomość potrzeby inwestowania w kadry, stosowane strategie rozwoju kadr).

Badania osób dorosłych:

- określenie dostępności pracowników w określonych zawodach,
- określenie dostępności i poziomu posiadanych kompetencji,
- zidentyfikowanie podejmowanych działań rozwojowych.

Badania instytucji i firm świadczących usługi rozwojowe:

- określenie charakterystyki instytucji i firm rozwojowych,
- określenie charakterystyki oferowanych form rozwoju kompetencji.

Tabela 1. W tabeli 1 przedstawiono poszczególne moduły badań w ujęciu czasowym.

Moduły badań	Edycja 2017	Edycja 2018	Edycja 2019	Edycja 2020	Edycja 2021	Edycja 2022
Pracodawcy – Panel						
Pracodawcy – Badanie przekrojowe						
Osoby dorosłe – Panel						
Osoby dorosłe – Badanie przekrojowe						
Usługi rozwojowe – Badanie przekrojowe						

Organizacja raportu

W raporcie opisano szczegółowo założenia metodologiczne oraz realizację badania Bilans Kapitału Ludzkiego 2016–2023¹ w trzech modułach:

1. Badania pracodawców.
2. Badania osób dorosłych.
3. Badania instytucji i firm świadczących usługi rozwojowe.

Opisy modułów badawczych składają się z następujących elementów:

- **Podstawowe założenia** – opis koncepcji badania wraz z tabelarycznym podsumowaniem realizacji kolejnych edycji.
- **Populacja** – opis populacji badanej, której dotyczą wnioski z badań.
- **Operat** – opis źródła danych, z którego wylosowano podmioty do badania.
- **Dobór próby** – ogólny opis koncepcji doboru próby.
- **Wagi** – opis wag analitycznych oraz warunków ich zastosowania.

Kwestionariusze zastosowane w badaniach terenowych są dostępne na stronie internetowej projektu. Zgodnie z koncepcją, po przeprowadzeniu badań terenowych w każdym kolejnym roku, trzon raportu uzupełniany będzie o część podsumowującą realizację.

¹ W przypadku szczegółów dotyczących realizacji badania wykorzystano zawartość raportów metodologicznych przesyłanych wraz z danymi przez Fundację Centrum Badania Opinii Społecznej (badania osób dorosłych) oraz PBS Sp. z o.o. (pozostałe moduły).

Badania pracodawców

Podstawowe założenia

Badania pracodawców zostaną zrealizowane w okresie 2017–2022 w dwóch formach: badań głównych oraz badań śródkresowych. Badania główne, realizowane co dwa lata, obejmują panel małych przedsiębiorstw (N = 1000), panel średnich przedsiębiorstw (N = 500), panel dużych przedsiębiorstw (N = 500) oraz próbę przekrojową (N = 1500) złożoną w równych częściach z małych, średnich i dużych przedsiębiorstw. Badania śródkresowe obejmą panel dużych przedsiębiorstw oraz panel średnich przedsiębiorstw.

Tabela 2. Zakładana liczebność próby w poszczególnych edycjach² badania

Badanie	Edycja 2017	Edycja 2018	Edycja 2019	Edycja 2020	Edycja 2021	Edycja 2022
Panel	2000	1000	2000	1000	2000	1000
Badanie przekrojowe	1500		1500		1500	

Tabela 3. Podsumowanie realizacji badań pracodawców – edycja 2017

a) Podstawowe informacje

Termin realizacji	21.08.2017–1.12.2017 (kontrola jakości 27.02.2018–06.04.2018)
Operat losowania	Polska Baza Danych Bisnode
Schemat losowania	warstwowy nieproporcjonalny
Technika zbierania danych	3481 wywiadów CAPI (96%) + 163 wywiady CATI (4%)
Liczba zrealizowanych wywiadów	3644
Średni czas trwania wywiadu	40 minut
Poziom netto realizacji próby	36%

² Kolorem ciemno szarym oznaczono edycje badania, a kolorem żółtym edycje śródkresowe.

b) Dobór próby

Aktualność operatu	06.2017
Zmienne warstwujące	wielkość, PKD
Liczebność wylosowanej próby (w tym zasadniczej)	11667 (9723)
Liczba wiązek	–
Liczebność wiązki	–

c) Realizacja

Liczebność próby wykorzystanej	11619
Kontakty błędne ³	900
Odmowy i inne przyczyny braku realizacji	7075

Populacja

Podobnie jak w Bilansie Kapitału Ludzkiego zrealizowanym w latach 2010–2015, badanie pracodawców obejmuje podmioty gospodarcze, które zatrudniają co najmniej jednego pracownika. Ze względu na przedmiot badania, jakim jest kapitał ludzki, z operatu wyłączone są:

1. Jednostki lokalne – wszelkiego rodzaju oddziały lub filie miejscowe podmiotów. W operacie będą znajdować się tylko „jednostki matki”, oznacza to, że każdy podmiot unikalnie identyfikował będzie numer REGON.
2. Podmioty z działu 94 działalność organizacji członkowskich według PKD 2007.
3. Podmioty z sekcji:
 - A – rolnictwo, leśnictwo, łowiectwo i rybactwo;
 - O – administracja publiczna i obrona narodowa; obowiązkowe zabezpieczenia społeczne;
 - T – gospodarstwa domowe zatrudniające pracowników; gospodarstwa domowe produkujące wyroby i świadczące usługi na własne potrzeby;
 - U – organizacje i zespoły eksterytorialne.
4. Podmioty o szczególnej formie prawnej:
 - 48 – fundacje;
 - 50 – Kościół Katolicki;
 - 51 – inne kościoły i związki wyznaniowe;
 - 55 – stowarzyszenia;

³ Opisane szczegółowo w części poświęconej doborowi próby.

- 60 – organizacje społeczne oddzielnie niewymienione;
- 70 – partie polityczne;
- 72 – związki zawodowe;
- 73 – organizacje pracodawców;
- 76 – samorząd gospodarczy i zawodowy;
- 85 – wspólnoty mieszkaniowe;
- 90 – związki grup producentów rolnych.

5. Podmioty samozatrudnione.

6. Podmioty, dla których w operacie niedostępne są dane teled adresowe niezbędne do przeprowadzenia badania.

Operat

Operat losowania stanowią podmioty znajdujące się w Polskiej Bazie Danych, będącej w posiadaniu firmy Bisnode. Baza ta została oparta na danych zawartych w Krajowym Rejestrze Urzędowym Podmiotów Gospodarki Narodowej (tzw. REGON). Dodatkowo zweryfikowano ją w wymiarze aktualności danych oraz uzupełniono o informacje teled adresowe dostępne w Internecie.

Dobór próby

Wylosowana liczebność próby uwzględnia zakładany poziom netto realizacji próby (36%), błąd operatu oraz kontakty rezerwowe. Zgodnie z wcześniejszymi doświadczeniami założono, że błąd operatu wraz z zapasowymi kontaktami to 16,7% wylosowanej próby.

Do obliczenia liczebności losowanej próby w edycji e badania (N_e^l) stosuje się następujący wzór:

$$N_e^l = 1,2 (N_e^z / 0,36)$$

gdzie:

1,2 – poprawka na błąd operatu i kontakty rezerwowe ($1/(1 - 0,167)$),

N_e^z – zakładana liczebność zrealizowanej próby w edycji e badania,

0,36 – zakładany poziom netto realizacji próby.

Próba badana podzielona jest na część zasadniczą oraz rezerwową. Próba rezerwowa to kontakty wykorzystywane do zastąpienia kontaktów błędnych, do których należą:

- niespełnienie przez podmiot kryteriów badania – podmiot nie spełnia kryteriów liczby zatrudnionych pracowników, wytypowanych do badania form prawnych lub sekcji PKD;
- błędne dane teleadresowe - prawidłowe dane nie są możliwe do odnalezienia przy wykorzystaniu źródeł ogólnodostępnych;
- podmioty nieistniejące – podmiot nigdy nie istniał lub obecnie jest w stanie zawieszenia albo likwidacji;
- duplikat rekordu.

Badania główne

Próba w badaniach głównych jest losowana z zastosowaniem warstwowania nieproporcjonalnego pod względem klasy wielkości podmiotu (3 klasy przedsiębiorstw):

- mikro i małe (zatrudniające od 2 do 49 pracowników),
- średnie (zatrudniające od 50 do 249 pracowników),
- duże (zatrudniające 250 lub większą liczbę pracowników)

oraz grup sekcji PKD, reprezentowanych w próbie równolicznie (6 kategorii):

1. Przemysł i górnictwo (sekcje B, C, D, E).
2. Budownictwo i transport (sekcje F, H).
3. Handel, zakwaterowanie, gastronomia, usługi wspierające (sekcje G, I, N).
4. Usługi specjalistyczne (sekcje J, K, L, M, R, S).
5. Edukacja (sekcja P).
6. Opieka zdrowotna i pomoc społeczna (sekcja Q).

Próba panelowa

Część badania pracodawców zrealizowana zostanie w formie panelowej. Średnie (N = 500) i duże (N = 500) podmioty będą badane w ramach panelu corocznie, tzn. zarówno w badaniach głównych, jak i badaniach śródk okresowych. Natomiast podmioty małe, zatrudniające mniej niż 50 pracowników, będą badane w liczebności 1000 co dwa lata, jedynie w badaniach głównych.

Próba przekrojowa

Próba przekrojowa w badaniach głównych (N = 1500) będzie równolicznie podzielona ze względu na klasę wielkości podmiotu. Tabela 4 zawiera dokładne informacje o założonych w kolejnych edycjach liczebnościach w poszczególnych klasach wielkości podmiotu.

Badania śródkresowe

Ze względu na małą dynamikę zjawisk będących przedmiotem badania w mikro i małych przedsiębiorstwach, badania śródkresowe realizowane w dwuletnich odstępach, między edycjami głównymi, obejmą swoim zakresem jedynie panel średnich i dużych przedsiębiorstw.

Osoba badana

W każdym z wylosowanych podmiotów osobą badaną udzielającą odpowiedzi jest osoba najlepiej zorientowana w kwestiach zatrudnienia, wymaganych umiejętności, rozwijania umiejętności pracowników (szkoleń). W zależności od wielkości i struktury podmiotów osobą kontaktową może być ktoś inny.

Wagi

Skonstruowano cztery rodzaje wag:

- 1) **trunc_dweight_smpl** – waga odzwierciedlająca strukturę populacji ze względu na zmienne warstwujące zastosowane na etapie doboru próby, w której obserwacje reprezentujące ekstremalnie wysoką liczbę podmiotów otrzymały wagę równą jeden. Jako kryterium ekstremalności przyjęto górny tzw. „zawias Tukey’a”, czyli wartość równą górnemu kwartylowi powiększonemu o dwukrotność rozstępu ćwiartkowego. Wagę należy stosować w przypadku wykonywania standardowych analiz. Przycięcia wag dokonywano w obrębie klas wielkości podmiotów;
- 2) **trunc_dweight_pop** – waga o strukturze **trunc_dweight_smpl**. Po jej zastosowaniu liczebność próby sumuje się do liczebności populacji. Wagę należy stosować w przypadkach, w których analizie poddane są zmienne mające wartościową interpretację w kategoriach populacyjnych, np. w celu oszacowania liczby pracowników, którzy są potrzebni w poszczególnych sektorach gospodarki;
- 3) **dweight_smpl** – waga analityczna będąca połączeniem wag związanych ze schematem doboru próby oraz wag non-response, która nadaje próbie strukturę populacyjną, jednak sumuje się do liczebności próby. Wagę stosuje się do kontroli efektu przycięcia zastosowanego we wcześniej opisanych wagach;
- 4) **dweight_pop** – waga o strukturze i zastosowaniu **dweight_smpl**. Po jej zastosowaniu liczebność próby sumuje się do liczebności populacji.

Tabela 4. Zakładane liczebności zrealizowanych prób w poszczególnych edycjach badania ze względu na uczestnictwo w panelu

a) Badanie główne: edycje 2017, 2019, 2021

Suma (edycja) = 3500

Wielkość podmiotu	2-49	50-249	250+
Panel (badania śródkresowe)	1000	500	500
Badanie przekrojowe (badanie główne)	500	500	500
Suma (edycja + wielkość)	1500	1000	1000

b) Badania śródkresowe: edycje 2018, 2020, 2022

Suma (edycja) = 1000

Wielkość podmiotu	2-49	50-249	250+
Panel (badania śródkresowe)	0	500	500
Badanie przekrojowe (badanie główne)	0	0	0
Suma (edycja + wielkość)	0	500	500

c) Suma

Suma dla wszystkich edycji = 13500

Wielkość podmiotu	2-49	50-249	250+
Panel (badania śródkresowe)	3000	3000	3000
Badanie przekrojowe (badanie główne)	1500	1500	1500
Suma (edycja + wielkość)	4500	4500	4500

Badania osób dorosłych

Podstawowe założenia

Badania osób zostaną zrealizowane w okresie 2017–2023 w dwóch formach: badań głównych oraz badań śródkresowych. Obydwie formy badań zostaną przeprowadzone trzykrotnie w dwuletnich odstępach czasu. Badania główne obejmą w każdej edycji po 4000 wywiadów, podzielonych na próbę panelową (N = 1500) oraz próbę przekrojową (N = 2500). Natomiast badania śródkresowe obejmą jedynie próbę panelową (N = 1500).

Tabela 5. Zakładana liczebność próby w poszczególnych edycjach⁴ badania

Badanie	Edycja 2017	Edycja 2018	Edycja 2019	Edycja 2020	Edycja 2021	Edycja 2022
Panel	1500	1000	1500	1500	1500	1500
Badanie przekrojowe	2500		2500		2500	

Tabela 6. Podsumowanie realizacji badań osób dorosłych – edycja 2017

a) Podstawowe informacje

Termin realizacji	21.08.2017–07.11.2017
Operat losowania	PESEL
Schemat losowania	warstwowy nieproporcjonalny
Technika zbierania danych	CAPI
Liczba zrealizowanych wywiadów	4056
Średni czas trwania wywiadu	37 minut
Poziom netto realizacji próby	56%

⁴ Kolorem ciemno szarym oznaczono edycje badania, a kolorem żółtym edycje śródkresowe.

b) Dobór próby

Aktualność operatu	07.2017
Zmienne warstwujące	podregion, wielkość miejscowości, płeć + wiek
Liczebność wylosowanej próby (w tym zasadniczej)	11609 (7144)
Liczba wiązek	893
Liczebność wiązki	13

c) Realizacja

Liczebność próby wykorzystanej	11319
Kontakty błędne ⁵	3552
Odmowy i inne przyczyny braku realizacji	3711

Populacja

Badaną populacją są polscy obywatele i obywatelki w wieku 18–69 lat, zamieszkujący na terytorium kraju.

Operat

Dane kontaktowe pochodzą z rejestru Polski Elektroniczny System Ewidencji Ludności (PESEL), zarządzanego przez Ministerstwo Cyfryzacji.

Dobór próby

Badania główne

Losowanie ma charakter warstwowy proporcjonalny ze względu na podregion GUS, klasę wielkości miejscowości oraz płeć w połączeniu z wiekiem.

W losowaniu stosuje się warstwowanie terytorialne obejmujące podział na podregiony GUS (NTS–3, tzn. 73 podregiony w skali kraju, po kilka w każdym województwie⁶⁶) oraz klasy wielkości miejscowości według GUS (9 klas):

⁵ Opisane szczegółowo w części poświęconej doborowi próby.

⁶ Zastosowano podział obowiązujący od 1 stycznia 2018 roku.

- wsie (tereny niezurbanizowane);
- miasta (tereny zurbanizowane), podzielone na klasy wielkości:
 - do 9999 mieszkańców;
 - 10 000–19 999 mieszkańców;
 - 20 000–49 999 mieszkańców;
 - 50 000–99 999 mieszkańców;
 - 100 000–199 999 mieszkańców;
 - 200 000–499 999 mieszkańców;
 - 500 000–999 999 mieszkańców;
 - Warszawa.

Opracowując założenia do edycji badania w 2017 r., wyróżniono 310 warstw wyznaczonych przez połączenie obu zmiennych. W ramach każdej klasy wielkości miejscowości dobierane są pierwotne jednostki losowania (PJL), tzn. gminy lub ich części traktowane jako osobne „jednostki administracyjne”. W przypadku gmin miejsko-wiejskich odrębnie traktowana jest część wiejska gminy (włączona do klasy 1 – wsie), a odrębnie część miejska gminy (miasto) – włączona do klasy miast, zgodnie z faktyczną liczbą mieszkańców. PJL są więc gmina wiejska, część wiejska gminy miejsko-wiejskiej, część miejska gminy miejsko-wiejskiej oraz gmina miejska. W pierwszym etapie losuje się PJL ze zwracaniem, proporcjonalnie do liczebności populacji w wieku 18–69, przy jednoczesnych założeniach, że wiązka realizacyjna wynosi trzynastcie osób oraz że każdą z warstw reprezentuje przynajmniej jedna wiązka.

W drugim etapie losowania, w ramach każdej PJL, stosuje się losowanie proste osób wchodzących w skład populacji badanej. Warstwowanie wylosowanych w drugim etapie osób obejmuje rozkład płci – kohort wiekowych dla poszczególnych klas wielkości miejscowości.

Jako kohorty wiekowe przyjmuje się następujące przedziały wieku:

- kobiety 18–29 lat;
- kobiety 30–39 lat;
- kobiety 40–49 lat;
- kobiety 50–69 lat;
- mężczyźni 18–29 lat;
- mężczyźni 30–39 lat;
- mężczyźni 40–49 lat;
- mężczyźni 50–69 lat.

Wylosowana liczebność uwzględnia zakładany poziom netto realizacji próby (56%), błąd operatu oraz kontakty rezerwowe. Zgodnie z wcześniejszymi doświadczeniami założono, że błąd operatu wraz z zapasowymi kontaktami to w przybliżeniu 38,5% wylosowanej próby. Do obliczenia liczebności losowanej próby w edycji e badania (N_e^l) stosuje się następujący wzór:

$$N_e^l = 1,625 (N_e^z / 0,56)$$

gdzie:

1,625 – poprawka na błąd operatu i kontakty rezerwowe ($1/(1 - 0,3846)$),

N_e^z – zakładana liczebność zrealizowanej próby w edycji e badania,

0,56 – zakładany poziom netto realizacji próby.

Liczebność wylosowanej próby zaokrąglana jest do wielokrotności liczby kontaktów w wiązce realizacyjnej.

Wiązki w badaniu są podzielone na część zasadniczą oraz rezerwową. Próba zasadnicza to ośmioosobowa część wiązki realizacyjnej złożona z kontaktów pozwalających na realizację badania przy założeniu przyjętego wskaźnika realizacji próby. Próba rezerwowa to pięcioosobowa część wiązki realizacyjnej obejmująca kontakty wykorzystywane do zastąpienia kontaktów błędnych, do których należą:

- trwała emigracja respondenta – osoba wylosowana do udziału w badaniu przebywa poza granicami kraju, a jej zamiary powrotu nie są znane;
- brak respondenta w miejscu zamieszkania – osoba wylosowana do udziału w badaniu (1) nie mieszka pod wskazanym adresem i nie jest znana osobom tam mieszkającym, lub (2) przebywa w gospodarstwie zbiorowym (zakładzie karnym, szpitalu etc.);
- nieprawidłowy adres – adres wskazany w operacie (1) nie istnieje, (2) jest pustostanem lub (3) nie jest mieszkaniem prywatnym;
- nieprawidłowe dane o respondencie – osoba wylosowana do badania (1) nie ukończyła 18. roku życia, (2) nie mówi po polsku lub (3) zmarła.

Badania śródkresowe/panelowe

Założycielska 1500-osobowa pod-próba panelowa została wylosowana spośród osób, które uczestniczyły w pierwszej edycji badania oraz wyraziły zgodę na ponowny kontakt. Dodatkowo jedynie osoby w wieku od 18 do 63 lat stanowiły operat losowania po edycji badania w 2017 r. Jest to zgodne z założeniem, że w ostatnim roku badania respondenci i respondentki uczestniczący w panelu nadal będą należeli/należały do populacji badanej, więc nie będą

starsze niż 69 lat. Struktura panelu, podobnie jak struktura badań przekrojowych, odzwierciedla strukturę próby w badaniach głównych. Respondenci wylosowani do uczestnictwa panelu, którzy z różnych przyczyn nie biorą udziału w kolejnych edycjach badania są zastępowani innymi osobami pochodzącymi z tej samej z warstwy. W celu gromadzenia jak największej liczby danych pozwalających na śledzenie losów osób badanych, osoby zastępujące są losowane w pierwszej kolejności spośród osób, które uczestniczyły w badaniu w edycjach wcześniejszych. Oznacza to także, że – z uwagi na krótki okres trwania panelu – w kolejnych latach nie są do niego dolosowywane osoby osiemnastoletnie.

Wagi

Skonstruowano cztery rodzaje wag:

- 1) **trunc_dweight_smpl** – waga odzwierciedlająca strukturę populacji ze względu na zastosowane zmienne warstwujące zastosowane na etapie doboru próby⁷, w której badani reprezentujący ekstremalnie wysoką liczbę osób otrzymały wagę równą jeden. Jako kryterium ekstremalności przyjęto górny tzw. „zawias Tukey’a”, czyli wartość równą górnemu kwartyłowi powiększonemu o dwukrotność rozstępu ćwiartkowego. Wagę należy stosować w przypadku wykonywania standardowych analiz;
- 2) **trunc_dweight_pop** – waga o strukturze trunc_dweight_smpl. Po jej zastosowaniu liczebność próby sumuje się do liczebności populacji. Wagę należy stosować w przypadkach, w których analizie poddane są zmienne mające wartościową interpretację w kategoriach populacyjnych, np. w celu oszacowania liczby osób w Polsce, które w minionych 12 miesiącach były zatrudnione na podstawie umów cywilno-prawnych;
- 3) **dweight_smpl** – waga analityczna będąca połączeniem wag związanych ze schematem doboru próby oraz wag non-response, która nadaje próbie strukturę populacyjną, jednak sumuje się do liczebności próby. Wagę stosuje się do kontroli efektu przycięcia zastosowanego we wcześniej opisanych wagach;
- 4) **dweight_pop** – waga o strukturze i zastosowaniu dweight_smpl. Po jej zastosowaniu liczebność próby sumuje się do liczebności populacji.

⁷ Z uwagi na stosunkowo mało liczną próbę, schemat ważenia nie odzwierciedla idealnie schematu losowania w przypadku podziału na podregiony oraz podziału na klasy wielkości miejscowości. Do ważenia zastosowano podział na 7 makroregionów (zgodnie z NTS 1) oraz podział na 4 klasy wielkości miejscowości, to jest (1) wsie, (2) małe miasta (< 20 tys. mieszkańców), (3) średnie miasta (20–99 tys. mieszkańców) oraz (4) duże miasta (> 100 tys. mieszkańców). W przypadku podziału na połączenie płci i wieku schemat ważenia odzwierciedla schemat doboru próby.

Badania instytucji i firm świadczących usługi rozwojowe

Podstawowe założenia

Badania osób zostaną zrealizowane w okresie 2017-2023 w formie badań głównych, które zostaną przeprowadzone trzykrotnie w dwuletnich odstępach czasu. Obejmą one w każdej edycji po 1000 wywiadów.

Tabela 7. Zakładana liczebność próby w poszczególnych edycjach badania

Badania	Edycja 2017	Edycja 2018	Edycja 2019	Edycja 2020	Edycja 2021	Edycja 2022
Próba przekrojowa	1000		1000		1000	

Tabela 8. Podsumowanie realizacji badań instytucji i firm świadczących usługi rozwojowe – edycja 2017

a) Podstawowe informacje

Termin realizacji	24.08.2017–07.12.2017 (kontrola jakości 26.02.2018–18.04.2018)
Operat losowania	baza własna
Schemat losowania	warstwowy nieproporcjonalny
Technika zbierania danych	859 wywiadów CAPI (86%) + 145 wywiadów CATI (14%)
Liczba zrealizowanych wywiadów	1004
Średni czas trwania wywiadu	43 minut
Poziom netto realizacji próby	31%

b) Dobór próby

Aktualność operatu	08.2017
Zmienne warstwujące	typ instytucji
Liczebność wylosowanej próby (w tym zasadniczej)	3388 (3226)
Liczba wiązek	–
Liczebność wiązek	–

c) Realizacja

Liczebność próby wykorzystanej	3363
Kontakty błędne ⁸	252
Odmowy i inne przyczyny braku realizacji	2107

Populacja

W badaniu Bilansu Kapitału Ludzkiego 2016–2023 zdecydowano się na uściślenie, a tym samym zawężenie definicji podmiotu identyfikowanego jako instytucja lub firma świadcząca usługi rozwojowe. Instytucja lub firma tego rodzaju będzie w tym badaniu rozumiana jako podmiot świadczący usługi rozwojowe (tj. kwalifikacyjne kursy zawodowe i kursy umiejętności zawodowych, certyfikowane programy rozwojowe, szkolenia i treningi grupowe, szkolenia i treningi indywidualne, konferencje, seminaria, szkoły dla dorosłych, studia podyplomowe, wykłady, odczyty, prelekcje, konsulting, doradztwo, coaching, mentoring, warsztaty, instruktaże) w ramach edukacji pozaformalnej i formalnej, wychodzącej poza standardowy cykl kształcenia⁹, jeżeli:

1. świadczy usługi rozwojowe na terytorium Polski dla osób dorosłych (18+),
2. jest niezależną instytucją/firmą, w tym firmą jednoosobową, z wyłączeniem tych podmiotów, które świadczą usługi rozwojowe tylko dla jednego klienta instytucjonalnego,
3. jest siedzibą główną instytucji/firmy wielooddziałowej lub jest jednym z oddziałów większej instytucji/firmy oraz posiada odrębną osobowość prawną, tzn. posiada odmienny od siedziby głównej/centrali numer działalności REGON,

⁸ Opisane szczegółowo w części poświęconej doborowi próby.

⁹ Za standardowy cykl kształcenia uznajemy kształcenie na poziomie podstawowym, gimnazjalnym, ponadgimnazjalnym i wyższym uczniów i studentów (I, II i III stopnia), będących najczęściej w wieku 6–24 lat, podejmujących najczęściej w sposób nieprzerwany kolejne etapy edukacji, korzystających najczęściej z usług edukacyjnych świadczonych w trybie stacjonarnym.

4. jest działalnością prowadzoną w ramach franczyzy,
5. charakter świadczonych usług rozwojowych przez nią nie obejmuje wyłącznie usług z poniższej listy:
 - a. kursów bezpieczeństwa i higieny pracy (BHP),
 - b. kursów bezpieczeństwa przeciwpożarowego (ppoż.),
 - c. kursów prowadzących do uzyskania uprawnień dla kierowców (szkoły nauki jazdy),
 - d. aktywności hobbistycznych (np. sportowych, rekreacyjnych),
 - e. aktywności o charakterze akcydentalnym (np. szkół rodzenia),
 - f. aktywności kulturalnych (np. usług świadczonych przez biblioteki, domy kultury),
 - g. usług rozwojowych skierowanych do sektora rolniczego (rolników, nie przedsiębiorców z terenów wiejskich), np. ośrodki doradztwa rolniczego,
 - h. usług o charakterze outsourcingowym (np. biur rachunkowych, biur radców prawnych),
6. nie należy do publicznych służb zatrudnienia (np. urzędy pracy),
7. uzyskała minimum 50 tysięcy PLN obrotu ze świadczenia usług rozwojowych (z wyłączeniem usług wymienionych w punkcie 5) dla osób dorosłych według danych za rok kalendarzowy poprzedzający badanie.

Z uwagi na fakt, że głównym obszarem zainteresowania są jednostki stabilne i posiadające tradycję świadczenia usług szkoleniowych, w I edycji badania trzon populacji będą stanowić jednostki, które istnieją na rynku usług rozwojowych przynajmniej trzy lata. W kolejnych edycjach badania do populacji zostaną włączone bazy zawierające dane jednostek prowadzących działalność krócej niż 3 lata, przy czym ich udział w wylosowanej próbie nie przekroczy 10% (w zależności od dostępnych danych udział tych podmiotów w poszczególnych warstwach może się różnić).

Operat

Podstawą do stworzenia kompletnego operatu losowania były istniejące bazy danych, w tym:

- dane gromadzone w System Informacji Oświatowej (dalej SIO),
- dane gromadzone w systemie POL-ON (dalej POL-ON),
- dane gromadzone w Bazie Usług Rozwojowych (dalej BUR),
- dane gromadzone w Rejestr Instytucji Szkoleniowych (dalej RIS),
- dane zebrane na potrzeby realizacji badania instytucji i firm szkoleniowych w ramach BKL w latach 2010–2015 (dalej CEAPP).

Baza stanowiąca potencjalny operat, składająca się z 30 350 podmiotów, przed podjęciem prac związanych z losowaniem, została odpowiednio uzupełniona oraz zweryfikowana.

Procedura ta trwała od czerwca do sierpnia 2017 roku i składała się z dwóch równoległe przeprowadzanych etapów:

1. Weryfikacji telefonicznej przynależności do populacji badanej.

Wyniki weryfikacji telefonicznej zostały przedstawione w tabeli 9.

Tabela 9. Wyniki telefonicznej weryfikacji przynależności do populacji badanej

Weryfikacja	Źródło: SIO	Źródło: POL-ON	Źródło: BUR	Źródło: CEAPP	Źródło: RIS	Źródło: Suma
weryfikacja negatywna	378	41	521	1561	1364	3865
weryfikacja pozytywna	217	34	141	631	959	1982
nawiązano kontakt przy braku weryfikacji	1924	252	1386	7487	5251	16300
nie udało się nawiązać kontaktu z podmiotem	392	54	296	6350	1111	8203
Suma	2911	381	2344	16029	8685	30350

2. Uzupelnienia i weryfikacji internetowej aktualności danych zawartych w poszczególnych bazach.

W wyniku tej procedury uzupełnione i zweryfikowane zostały dane, które są konieczne do przeprowadzenia badania, tzn.:

- adres podmiotu,
- numer telefonu,
- numer REGON.

Dzięki uzupełnieniom możliwe było usunięcie z bazy podmiotów nieaktualnych oraz zduplikowanych w kilku bazach (5864 podmioty).

W kolejnych edycjach badania baza teleadresowa będzie weryfikowana i uzupełniana, mając na względzie zmiany na rynku instytucji i firm rozwojowych (zakończenie działalności niektórych instytucji, rozpoczęcie działalności innych instytucji, zmiany siedzib, adresów itp.). Każdego roku próba instytucji i firm świadczących usługi rozwojowe podzielona zostanie losowo na próbę zasadniczą oraz próbę rezerwową. Wielkość próby zasadniczej zostanie obliczona w ten sposób, aby można było zrealizować badanie, uwzględniając stopy zwrotu założone przez Wykonawcę.

Dobór próby

Z uwagi na fakt, że jednym z celów badania była możliwość dokonania porównań między różnymi rodzajami podmiotów, o ostatecznej przynależności do operatu badania zdecydowała przynależność do konkretnej warstwy losowania. Warstwowanie wyznaczone zostało przez typ podmiotu, zgodny z jego przynależnością do pierwotnego źródła danych.

Zidentyfikowano cztery typy podmiotów:

- 1) podmioty obecne w bazie SIO – CKU, CKP, ODZ, szkoły ponadgimnazjalne,
- 2) podmioty obecne w bazie POL-ON – uczelnie wyższe,
- 3) podmioty obecne w BUR – prywatne podmioty gospodarcze zarejestrowane w Bazie Usług Rozwojowych,
- 4) podmioty obecne w CEAPP lub RIS – prywatne podmioty gospodarcze niezarejestrowane w bazie BUR.

W tabeli 10 przedstawiono wyszczególnione warstwy oraz ich statusy weryfikacyjne zarówno w operacie losowania, jak i w wylosowanej próbie.

Tabela 10. Operat losowania oraz wylosowana próba według wyników weryfikacji telefonicznej

Weryfikacja	SIO	POL-ON	BUR	CEAPP	RIS	Suma
Operat: weryfikacja pozytywna	208	137	137	608	955	1942
Operat: nawiązano kontakt przy braku weryfikacji	1618	1288	1288	0	0	3121
Operat: nie udało się nawiązać kontaktu z podmiotem	148	170	170	0	0	370
Operat: Suma	1974	1595	1595	1563	5433	
Próba: weryfikacja pozytywna	208	137	137	413	616	1408
Próba: nawiązano kontakt przy braku weryfikacji	690	735	735	0	0	1640
Próba: nie udało się nawiązać kontaktu z podmiotem	131	157	157	0	0	340
Próba: Suma	1029	1029	1029	1029	3388	

Wylosowana liczebność próby uwzględnia zakładany poziom netto realizacji próby (31%), błąd operatu oraz kontakty rezerwowe. Z uwagi na wcześniejsze prace weryfikacyjne, założono, że błąd operatu wraz z zapasowymi kontaktami to 5% wylosowanej próby.

Do obliczenia liczebności losowanej próby w edycji e badania (N_e^l) stosuje się następujący wzór:

$$N_e^l = 1,05 (N_e^z / 0,31)$$

gdzie:

1,05 – poprawka na błąd operatu i kontakty rezerwowe,

N_e^z – zakładana liczebność zrealizowanej próby w edycji e badania,

0,31 – zakładany poziom netto realizacji próby.

Próba badana podzielona jest na część zasadniczą oraz rezerwową. Próba rezerwowa to kontakty wykorzystywane do zastąpienia kontaktów błędnych, do których należą:

- niespełnienie przez podmiot kryteriów badania – podmiot nie spełnia kryteriów liczby zatrudnionych pracowników, wytypowanych do badania form prawnych lub sekcji PKD;
- błędne dane teleadresowe – prawidłowe dane nie są możliwe do odnalezienia przy wykorzystaniu źródeł ogólnodostępnych;
- podmioty nieistniejące – podmiot nigdy nie istniał lub obecnie jest w stanie zawieszenia albo likwidacji.

Wagi

Z uwagi na to, że zastosowano w doborze próby oraz w ważeniu danych tą samą, pojedynczą zmienną warstwową, a tym samym w obrębie warstwy wszystkie obserwacje miały te same wagi analityczne, skonstruowano dwa rodzaje wag:

- 1) **dweight_smpl** – waga analityczna będąca połączeniem wag związanych ze schematem doboru próby oraz wag non-response, która nadaje próbie strukturę populacyjną, jednak sumuje się do liczebności próby. Wagę należy stosować w przypadku wykonywania standardowych analiz;
- 2) **dweight_pop** – waga o strukturze i zastosowaniu dweight_smpl. Po jej zastosowaniu liczebność próby sumuje się do liczebności populacji. Wagę należy stosować w przypadkach, w których analizie poddane są zmienne mające wartościową interpretację w kategoriach populacyjnych, np. w celu oszacowania liczby osób w Polsce, które w minionych 12 miesiącach uczestniczyły w różnego rodzaju szkoleniach.



www.parp.gov.pl

