

Monitorowanie Losów Edukacyjno- Zawodowych Absolwentów i Młodych Dorosłych - raport metodologiczny

Warszawa 2019

Autorzy:

Krzysztof Bulkowski

dr Paweł Grygiel

Grzegorz Humenny

dr Michał Sitek

Jędrzej Stasiowski

dr Tomasz Żółtak

© Copyright by: Instytut Badań Edukacyjnych, Warszawa, marzec 2019

Wzór cytowania:

Bulkowski, K., Grygiel, P., Humenny, G., Sitek, M., Stasiowski, J. i Żółtak, T. (2019). Monitorowanie Losów Edukacyjno-Zawodowych Absolwentów i Młodych Dorosłych – raport metodologiczny. Warszawa: Instytut Badań Edukacyjnych.

Wydawca:

Instytut Badań Edukacyjnych

ul. Górczewska 8

01-180 Warszawa

tel. (22) 241 71 00; www.ibe.edu.pl

Egzemplarz bezpłatny

Streszczenie

Raport zawiera opis założeń oraz wdrożonych rozwiązań metodologicznych w badaniach realizowanych w ramach projektu Monitorowanie Losów Edukacyjno-Zawodowych Absolwentów i Młodych Dorosłych. Szczegółowo opisano badania ilościowe przeprowadzone w ramach pierwszej rundy monitoringu realizowano za pomocą dwóch technik: ankiety audytoryjnej oraz wywiadów przeprowadzanych na podstawie standaryzowanego kwestionariusza. Badania zostały przeprowadzone z uczniami ostatnich klas zasadniczych szkół zawodowych, techników i szkół policealnych, a także z dyrektorami szkół zawodowych. Moduł badania uczniów za pomocą wywiadów miał charakter panelowy – do respondentów powrócono kilka miesięcy po zakończeniu szkoły. Kolejne rozdziały poświęcone są poszczególnym etapom procesu badawczego: celom badania, doborowi próby, konstrukcji narzędzi badawczych, realizacji terenowej badania, przetwarzaniu wyników badania, w tym szczegółom związanym z konstrukcją zbiorów danych i poszczególnych zmiennych, a także wskaźnikom prezentowanym w raportach dla szkół. Runda pierwsza zawierała również moduł jakościowy, w którym przeprowadzono zogniskowane wywiady grupowe oraz pogłębione wywiady indywidualne z dyrektorami, doradcami edukacyjno-zawodowymi, uczniami i absolwentami szkół zawodowych, nauczycielami przedmiotów zawodowych, a także z pracodawcami i przedstawicielami jednostek samorządu terytorialnego na poziomie powiatu. W raporcie opisane są również badania ilościowe przeprowadzone w ramach rundy pilotażowej, poprzedzającej pierwszą rundę monitoringu, dzięki którym wypracowano ostateczne rozwiązania metodologiczne zastosowane w projekcie. W aneksie do raportu zamieszczone jest szesnaście narzędzi zastosowanych w opisywanych badaniach: kwestionariusze użyte w badaniach ilościowych oraz scenariusze wywiadów, na podstawie których zrealizowano badanie jakościowe.

Spis Treści

Streszczenie	3
Spis Treści	4
1. Wprowadzenie	6
2. Cele i schemat badań.....	8
3. Dobór próby.....	10
3.1. Założenia podstawowe	10
3.2. Teoretyczne podstawy optymalnej konstrukcji próby do badania	10
3.3. Koncepcja doboru próby: badanie audytoryjne	12
3.4. Koncepcja doboru próby: badanie CAPI	16
3.5. Dobór próby a ważenie danych	19
3.6. Szacowanie błędów standardowych z uwzględnieniem schematu doboru próby	20
3.7. Operat losowania	21
4. Narzędzia badawcze.....	22
4.1. Kwestionariusz wywiadu z dyrektorem szkoły	23
4.2. Ankieta audytoryjna	24
4.3. Kwestionariusz do wywiadu z uczniem	25
4.4. Kwestionariusz do wywiadu z absolwentem	27
5. Realizacja badania w terenie	30
5.1. Organizacja realizacji badania	30
5.2. Techniki i techniki gromadzenia danych	31
5.3. Realizacja próby	33
5.4. Problemy w badaniu terenowym	35
6. Przetwarzanie danych	36
6.1. Struktura podstawowych zbiorów danych	36
6.2. Ważenie poststratyfikacyjne	43
6.3. Zbiory <i>epizodów</i>	48
6.4. Zbiory osobo-miesiący	61
7. Wskaźniki w raportach dla szkół.....	67
7.1. Procedura obliczania wskaźników	67
7.2. Wskaźniki obliczane na poziomie indywidualnym	68
7.3. Wskaźniki obliczane na poziomie zagregowanym	81
8. Charakterystyka respondentów	93

9. Runda pilotażowa	100
9.1. Rola Badania Losów Absolwentów Szkół Zawodowych (BLASZ) w projekcie MLEZAiMD	100
9.2. Narzędzia	101
9.3. Przyczyny uniemożliwiający uzyskanie wywiadów	102
9.4. Ograniczenia próby zrealizowanej w Badaniu Losów Absolwentów Szkół Zawodowych	103
9.5. Ograniczenia próby zrealizowanej w pilotażowej rundzie monitoringu	104
9.6. Możliwość uogólniania wniosków z analizy wyników pilotażowej rundy monitoringu	104
10. Badanie jakościowe	110
10.1. Narzędzia	110
10.2. Założenia doboru próby badawczej	112
10.3. Realizacja badania w terenie	113
10.4. Liczba zrealizowanych wywiadów	114
10.5. Czas trwania wywiadów	114
10.6. Trudności realizacyjne	114
Literatura cytowana	115
Aneks – narzędzia badawcze.....	115

1. Wprowadzenie

W ramach prac nad opracowaniem, przetestowaniem i przygotowaniem do wdrożenia systemowych narzędzi pozwalających na systematyczne śledzenie i monitorowanie losów zawodowych absolwentów szkół zawodowych w latach 2016-2019 Instytut Badań Edukacyjnych przeprowadził projekt „Monitorowanie Losów Edukacyjno-Zawodowych Absolwentów i Młodych Dorosłych” (MLEZAIMD, sygnatura: POWR.02.15.00-IP.02-00-004/16). Jednym z jego celów było stworzenie metodologii badań empirycznych, które wspomogłyby budowę systemu monitorowania losów absolwentów szkół zawodowych. Prezentowany raport zawiera opis założeń oraz wdrożonych rozwiązań metodologicznych, które zostały zastosowane w przeprowadzonych badaniach terenowych.

Projekt MLEZAIMD jest realizowany w ścisłym powiązaniu z projektem realizowanym w trybie konkursowym „Monitorowanie losów absolwentów szkół zawodowych - pierwsza runda” (MLASZ) w ramach działania „Zapewnienie informacji zwrotnej dla systemu kształcenia zawodowego na temat losów absolwentów szkół zawodowych (sygnatura POWR.02.15.00-IP.02-00-003/16)”, w ramach którego beneficjent (PBS Sp. z o.o.) miał za zadanie przeprowadzić badania terenowe. W projekcie konkursowym wykorzystywane były produkty (próby badawcze, szczegółowe instrukcje dotyczące rozwiązań metodologicznych oraz narzędzia badawcze: kwestionariusze i scenariusze wywiadów) wytworzone w ramach projektu MLEZAIMD. Natomiast wytworzone w ramach projektu konkursowego (MLASZ) produkty, czyli zbiory danych, transkrypcje wywiadów oraz informacje dotyczące realizacji badania są przedmiotem analiz w projekcie MLEZAIMD. Wyniki tych badań zawarte są w raporcie końcowym podsumowującym realizację projektu. Niniejszy dokument jest niejako raportem „technicznym” będącym dokumentacją założeń badawczych oraz pokazującym efekty badań terenowych z punktu widzenia tychże założeń.

Główne badania zaplanowane w ramach projektu MLEZAIMD, czyli pierwsza pełna runda monitoringu, były badaniami ilościowymi przeprowadzanymi wśród uczniów i absolwentów szkół zawodowych: zasadniczych szkół zawodowych, techników i szkół policealnych. Realizowano je za pomocą dwóch technik: ankiety audytoryjnej oraz wywiadów przeprowadzanych na podstawie standaryzowanego kwestionariusza. Badania zostały przeprowadzone z uczniami ostatnich klas zasadniczych szkół zawodowych, techników i szkół policealnych, a także z dyrektorami szkół zawodowych. Moduł badania uczniów za pomocą wywiadów miał charakter panelowy – do respondentów powrócono kilka miesięcy po zakończeniu szkoły. Kolejne rozdziały poświęcone są poszczególnym etapom procesu badawczego: celom badania (Rozdział 2), doborowi próby (Rozdział 3), konstrukcji narzędzi badawczych (Rozdział 4), realizacji terenowej badania (Rozdział 5), przetwarzaniu wyników badania, w tym szczegółom związanym z konstrukcją zbiorów danych i poszczególnych zmiennych (Rozdział 6), a także wskaźnikom prezentowanym w raportach dla szkół (Rozdział 7).

Pierwsza runda monitoringu zawierała również moduł jakościowy, w którym przeprowadzono zogniskowane wywiady grupowe oraz pogłębione wywiady indywidualne z dyrektorami, doradcami edukacyjno-zawodowymi, uczniami

i absolwentami szkół zawodowych, nauczycielami przedmiotów zawodowych, a także z pracodawcami i przedstawicielami jednostek samorządu terytorialnego na poziomie powiatu. Szczegóły metodologiczne dotyczące założeń realizacyjnych tego modułu zawarte są w Rozdziale 8.

Ostatni rozdział (9) dotyczy działań podjętych przed pierwszą pełną rundą monitoringu. Była to tzw. runda pilotażowa, w której wykorzystano wcześniejsze badanie przeprowadzone wśród uczniów szkół zawodowych. W ramach opisywanego tu projektu MLEZAiMD do uczniów tych powrócono po niecałych dwóch latach od planowego ukończenia szkoły. Były to badania, dzięki którym wypracowano rozwiązania metodologiczne i narzędzia do pierwszej pełnej rundy monitoringu. Analizę wyników rundy pilotażowej znaleźć można w raporcie „Losy edukacyjne i zawodowe absolwentów. Raport z pilotażowej rundy monitoringu losów zawodowych absolwentów szkół zawodowych”, który został opublikowany w grudniu 2017 roku.

Ze względu na „dokumentacyjny” charakter niniejszego raportu w aneksie zamieszczone są narzędzia badawcze, które zostały wypracowane w ramach projektu MLEZAiMD. Zawierają one dokładne treści pytań, instrukcji dla ankierów i reguł przejścia itd. w ostatecznych wersjach, które zostały wykorzystane w projekcie konkursowym w trakcie badań terenowych. Jest to szesnaście narzędzi w tym: 8 kwestionariuszy użytych w badaniach ilościowych oraz 8 scenariuszy wywiadów, na podstawie których zrealizowano badania jakościowe.

2. Cele i schemat badań

Nadrzędnym celem projektu MLEZAiMD jest opracowanie, przetestowanie i przygotowanie do wdrożenia systemowych narzędzi pozwalających na monitorowanie losów zawodowych absolwentów szkół zawodowych na poziomie krajowym, regionalnym i lokalnym. Narzędzia te mają wykorzystywać dostępne dane administracyjne, ograniczając w ten sposób konieczność prowadzenia kosztownych badań empirycznych i umożliwiając dostęp do informacji o możliwie dużej liczbie szkół. Opracowanie takich narzędzi wymaga zidentyfikowania potencjalnych źródeł danych, ich dostępności oraz możliwości ich wykorzystania do prowadzenia analiz, formułowania wniosków i rekomendacji. Głównym celem opisywanych w tym raporcie badań empirycznych jest zidentyfikowanie trafnych wskaźników, które mogłyby być podstawą systematycznego monitoringu. Zebrane dane, w szczególności sondażowe, posłużyły do zidentyfikowania zmiennych niezbędnych do realizacji tego celu.

Założono, że najbardziej efektywnym sposobem dotarcia do absolwentów szkół zawodowych jest „zrekrutowanie” ich w ostatniej klasie, na końcowym etapie ich kształcenia. Niewątpliwym atutem tego rozwiązania jest możliwość skorzystania z zawartych w Systemie Informacji Oświatowej informacji o szkołach pozwalających na losowy dobór reprezentatywnej próby uczniów.

W jednym z modułów badania ilościowego wykorzystano metodologię badania panelowego. W trakcie nauki szkolnej z uczniami przeprowadzany był indywidualny wywiad na podstawie specjalnie w tym celu przygotowanego kwestionariusza, a także zbierane dane kontaktowe i odpowiednie zgody na przetwarzanie danych, niezbędne do ponownego kontaktu po zakończeniu szkoły. W ramach tego drugiego kontaktu przeprowadzany był drugi wywiad, już z absolwentami, ponad 9 miesięcy po zakończeniu kształcenia w szkole. Wywiad prowadzony był na podstawie innego wytworzonego w projekcie narzędzia: kwestionariusza absolwenta szkoły zawodowej.

Drugi moduł badania ilościowego miał charakter badania przekrojowego. Uczniowie byli badani za pomocą krótkiej ankiety przeprowadzanej „audytoryjnie” we wszystkich oddziałach ostatniej klasy danej szkoły zawodowej. Oprócz tego uczniowie wypełniali również zgodę na przetwarzanie danych oraz zgodę na pozyskiwanie przez Instytut Badań Edukacyjnych danych dotyczących aktywności edukacyjnej i zawodowej ucznia, w tym danych osobowych, gromadzonych w bazach danych instytucji publicznych: Okręgowych Komisji Egzaminacyjnych, Centralnej Komisji Egzaminacyjnej, Zakładu Ubezpieczeń Społecznych, Ministerstwa Edukacji Narodowej oraz Ministerstwa Nauki i Szkolnictwa Wyższego. Celem tego drugiego modułu badania było zatem zgromadzenie reprezentatywnej i wystarczająco dużej próby uczniów, dzięki której można byłoby monitorować ich aktywność zawodową i edukacyjną wykorzystując dane rejestrowe.

W obu opisywanych modułach postanowiono zebrać również informacje o szkole, w której przeprowadzane było badanie. Do tego celu stworzono kwestionariusz do wywiadu z dyrektorem szkoły. Dzięki tym danym wyniki zebrane wśród uczniów można osadzić w szerszym kontekście warunków związanych ze szkołą, w której się kształca.

Budowanie systemu monitorowania losów absolwentów wymaga również zbadania potrzeb i oczekiwań interesariuszy – przyszłych użytkowników przygotowywanych rozwiązań. Było to głównym celem zaprojektowanych badania jakościowych. Zdecydowano się na wykorzystanie dwóch najbardziej popularnych metod: indywidualnego wywiadu pogłębionego (IDI) oraz zogniskowanego wywiadu grupowego (FGI). Respondentami w badaniu byli dyrektorzy, nauczyciele, doradcy zawodowi oraz uczniowie i absolwenci szkół zawodowych (ZSZ oraz techników), jak również pracodawcy współpracujący i niewspółpracujący ze szkołami zawodowymi oraz przedstawiciele samorządów. By uchwycić różnorodność opinii, potrzeb i oczekiwań interesariuszy, realizację badań zaprojektowano w dwóch regionach, które reprezentowały odmienne warunki na lokalnym rynku pracy, jak i w zakresie organizacji systemu kształcenia zawodowego. Jedną z lokalizacji miało być duże miasto z niskim poziomem bezrobocia i skoncentrowaną terytorialnie siecią szkół kształcących zawodowo, drugą natomiast bardziej rozległy region kilku przylegających do siebie powiatów, który charakteryzowałby się relatywnie trudną sytuacją na rynku pracy i rozproszoną siecią szkół zawodowych.

3. Dobór próby

3.1. Założenia podstawowe

Badanie obejmuje uczniów w wieku do 29 lat ostatnich klas/roku kształcenia w szkołach zawodowych (Zasadniczych Szkołach Zawodowych, Technikach oraz Szkołach Policealnych)¹.

Badanie, zgodnie z założeniami, ma składać się z dwóch części:

- Badania z użyciem ankiety audytoryjnej, realizowanej z wszystkimi uczniami ostatniej klasy/roku kształcenia. Badanie ma być zrealizowane w 750 szkołach; przewiduje się, że liczba uczniów przewidzianych do badania wyniesie około 30.000.
- Badania z wykorzystaniem wywiadów CAPI, realizowanych z próbą uczniów ostatniej klasy/roku kształcenia. Badanie ma być zrealizowane w 600 szkołach, zaś – przy przewidywanym przeciętnym poziomie realizacji próby około 65% - liczba przebadanych uczniów powinna wynieść około 10.000.

Próba – w obu częściach badania – powinna zapewnić reprezentatywność na poziomie obszarów kształcenia zawodowego i województw.

Opisana w tym rozdziale koncepcja doboru próby opisuje procedurę:

- doboru próby szkół,
- doboru próby szkół rezerwowych,
- doboru uczniów mających brać udział w badaniu CAPI.

Ze schematu doboru próby wynika również zasada tworzenia systemu wag posstratyfikacyjnych, a częściowo również postrealizacyjnych, które również są elementem przedstawionej niżej koncepcji.

3.2. Teoretyczne podstawy optymalnej konstrukcji próby do badania

Najwłaściwszym sposobem doboru próby reprezentatywnej jest dobór losowy. Próba losowa jest z definicji próbą reprezentatywną – pozwala uogólniać uzyskane wyniki zarówno na całą populację, jak i na jej części, zaś błędy oszacowań mają charakter

¹ Niniejszy rozdział zawiera ekspertyzę wytworzoną przed rozpoczęciem badań terenowych Haman J., *Metodologia doboru próby uczniów ostatnich klas szkół zawodowych w badaniu Monitorowanie Losów Edukacyjno-Zawodowych Absolwentów i Młodych Dorosłych*, dokument wewnętrzny projektu (15.12.2016).

błędów losowych i jakkolwiek nie dają się w konkretnym przypadku przewidzieć, to ich ogólne własności opisywane są zgodnie z regułami rachunku prawdopodobieństwa.

W węższym znaczeniu można powiedzieć, że próba losowa jest reprezentatywna na poziomie takiej części populacji, dla jakich przewidywane błędy losowe oszacowań nie przekraczają akceptowalnych poziomów. Wielkość oczekiwanych błędów losowych (operacjonalizowana zwykle jako „błąd standardowy”, lub dokładność oszacowania, czyli połowa szerokości 95% przedziału ufności, odpowiadająca 1,96 „błędowi standardowemu”) zależy z kolei od wielkości próby i sposobu jej doboru, z drugiej zaś strony – od niezależnych od konstrukcji próby charakterystyk mierzonych wartości (a więc – w przypadku szacowania średnich lub odsetków – od wielkości wariancji mierzonej zmiennej).

W przypadku badania MLEZaiMD część istotnych charakterystyk próby została przyjęta z góry – dotyczy to przede wszystkim jej liczebności. Charakter badania narzuca również dwustopniowy charakter próby do badania CAPI (losowanie szkół – losowanie uczniów) oraz zespołowy charakter próby do badania audytoryjnego (losowanie szkół – badanie wszystkich spełniających założone warunki uczniów w szkole). Ze względu na duże różnice między poszczególnymi typami badanych szkół, naturalne jest także traktowanie każdego z nich jako warstwy, z której doboru dokonuje się niezależnie.

Metodą doboru próby szkół powszechnie stosowaną w badaniach edukacyjnych jest losowanie systematyczne z wykorzystaniem warstw explicite oraz implicate. Jest to rozwiązanie pozwalające łatwo realizować schematy warstwowe uwzględniające hierarchicznie wiele zmiennych warstwowujących, a także dość odporne na możliwe błędy, np. związane z błędami w operacie (dane służące do przypisania jednostki do warstwy jednocześnie wpływają na liczbę jednostek alokowanych do warstwy).

W tej sytuacji pozostaje do określenia:

1. Wyznaczenie liczby warstw explicite oraz określenie liczby szkół przydzielanych do każdej z warstw.
2. Określenie systemu warstw implicate (a więc zasady porządkowania operatu).
3. Określenie sposobu wyznaczania wag poszczególnych szkół (tj. prawdopodobieństw wylosowania).
4. Określenie sposobu doboru szkół rezerwowych
5. Określenie technicznej procedury losowania szkół:
6. Opis procedury doboru uczniów:
 - a. do badania audytoryjnego;
 - b. do badania CAPI.

3.3. Koncepcja doboru próby: badanie audytoryjne

3.3.1. Warstwy explicite

Jako warstwy explicite należy przyjąć:

- warstwę średnich szkół zawodowych (techników)
- warstwę zasadniczych szkół zawodowych
- warstwę szkół policealnych.

W roku szkolnym 2016/17 nie działają już szkoły zasadnicze oraz technika dla dorosłych. Szkoły policealne formalnie podzielone są na „szkoły dla młodzieży” oraz „szkoły dla dorosłych”, ponieważ jednak w praktyce i jedne i drugie kierują swoją ofertę do osób dorosłych one do osób, nie ma podstaw, do rozdzielenia ich w osobne warstwy.

Ponieważ warstwy traktujemy osobno, i w dużej mierze osobno prowadzona będzie analiza danych uzyskanych w badaniu, kwestia określenia liczby szkół dla poszczególnych warstw rozstrzygnięta może być na zasadzie arbitralnej decyzji, biorącej pod uwagę:

- potrzeby badawcze (traktowane łącznie z badaniem CAPI)
- możliwości i ograniczenia organizacyjne, w tym również ryzyka związane z badaniami w poszczególnych warstwach.

Po analizie potrzeb, ograniczeń i ryzyk zaproponowano następujące liczebności szkół w warstwach:

- warstwa średnich szkół zawodowych (techników): 150 szkół
- warstwa zasadniczych szkół zawodowych: 300 szkół
- warstwa szkół policealnych: 300 szkół.

3.3.2. Warstwy implicite

Dla optymalizacji reprezentatywności doboru szkół, szczególnie ze względu na obszar kształcenia oraz województwo, zaproponowano przyjęcie następujących cech szkoły jako zmiennych, ze względu na które nastąpi hierarchiczne porządkowanie zbioru danych do losowania systematycznego, a więc jako warstwy implicite:

- Dominujący obszar kształcenia. Jako dominujący obszar kształcenia rozumie się ten obszar kształcenia, do którego przypisana jest największa część uczniów szkoły, wedle danych SIO z bieżącego roku szkolnego. Każdej szkole przypisany jest dokładnie jeden dominujący obszar kształcenia, także wtedy, gdy szkoła kształci również w innych obszarach. Przypisanie szkoły do danego obszaru kształcenia nie oznacza jednak, że – jeśli szkoła zostanie wylosowana do badania – to jedynie uczniowie z danego obszaru kształcenia będą uczestniczyli

w badaniu. Przypisanie szkoły do obszaru kształcenia ma charakter jedynie techniczno-informacyjny.

- Województwo
- Klasa wielkości miejscowości
- Szacowana liczba uczniów nie starszych niż 29 lat w ostatniej klasie/ostatnim roku nauki. Wartość tę wyznaczamy – na podstawie danych SIO – jako liczbę uczniów w ostatniej klasie, przemnożoną przez proporcję liczby uczniów szkoły urodzonych nie wcześniej niż w 1988 roku do liczby wszystkich uczniów szkoły.

3.3.3. Wagi (prawdopodobieństwa wylosowania) poszczególnych szkół

Ponieważ w badaniu audytoryjnym mają uczestniczyć wszyscy uczniowie szkoły kwalifikujący się do badania, równość prawdopodobieństwa trafienia do próby dla wszystkich jednostek z badanej populacji osiągnięta by została przy zastosowaniu doboru szkół z równymi prawdopodobieństwami – a zatem przy zastosowaniu dla wszystkich szkół takiej samej wagi (równej 1). Ze względu na optymalizację kosztów i uproszczenie organizacji badania zaproponowano wprowadzenie do tej reguły dwóch modyfikacji:

- Z badania wyłączone szkoły o przewidywanej liczbie uczniów do przebadania mniejszej niż 10. Przy bardzo nieznacznym ograniczeniu reprezentatywności badania (poprzez niepełne pokrycie próbą populacji – wykluczamy z badania ok. 2,7% ogółu populacji), pozwoli to na optymalizację kosztów badania (przy dużym udziale stałych kosztów dotarcia do szkoły, badanie ucznia szkoły bardzo małej byłoby nieproporcjonalnie drogie).
- Dla szkół liczących ponad 100 uczniów przewidzianych do badania przyjęto jako wagę losowania wartość $(100/[\text{liczba uczniów kwalifikujących się do badania}])$; a więc zmniejszono prawdopodobieństwo ich wylosowania. W tym przypadku decydują o tym względy logistyczne: szkoły duże, a zwłaszcza (nieliczne) szkoły bardzo duże generować będą w badaniu szczególnie wysokie koszty; jednocześnie należy się liczyć z ich ponadprzeciętną niechęcią do kooperowania z badaczami.
- Aby zmniejszenie prawdopodobieństwa ich doboru nie prowadziło do obciążenia uzyskanych w badaniu wyników, uczniowie pochodzący z tych szkół uzyskają w analizie danych dodatkową wagę korygującą odpowiadającą odwrotności wagi losowania, czyli $([\text{liczba uczniów kwalifikujących się do badania}]/100)$.

3.3.4. Dobór szkół rezerwowych

Standardowym rozwiązaniem stosowanym w badaniach edukacyjnych jest dobór dla każdej szkoły z próby zasadniczej 2 dwóch szkół rezerwowych: szkół znajdujących się na uporządkowanym operacie losowania bezpośrednio za i przed daną szkołą wylosowaną do badania. Przy zastosowaniu tej zasady, zbiór szkół rezerwowych jest ograniczony, zaś każda szkoła w nim się znajdująca pełni może służyć jako rezerwa tylko dla jednej szkoły

z próby zasadniczej. W przypadku badania MLEZAiMD zastosowanie tej procedury mogłoby być utrudnione (zwłaszcza w module CAPI), ze względu na liczbę szkół biorących udział w badaniu. Proponuję zatem pewną modyfikację tej reguły, pozwalającą na traktowanie wszystkich szkół w kraju jako szkół rezerwowych, przy zachowaniu zasady doboru jako szkoły rezerwowej szkoły najbardziej podobnej do szkoły wylosowanej do próby zasadniczej, oraz zasady jednoznaczności wskazywania szkół rezerwowych.

Zgodnie z tą propozycją:

Każda szkoła kwalifikująca się do udziału w badaniu ma przypisaną „pozycję w bazie”: określającą pozycję w uporządkowanym operacie do losowania systematycznego. W konsekwencji „pozycja w bazie” określona jest zarówno dla szkół wylosowanych do próby zasadniczej, jak i dla każdej szkoły nie wylosowanej do próby (a więc - potencjalnej szkoły rezerwowej).

Szkołę zastępczą wykorzystujemy, gdy:

- Wylosowana szkoła (lub szkoła wcześniej wskazana jako zastępcza) odmawia udziału w badaniu.
- Wylosowana szkoła (lub szkoła wskazana jako zastępcza) nie ma żadnych uczniów ostatniej klasy urodzonych w latach 1988-2001.
- W wylosowanej szkole (lub szkole wskazanej jako zastępcza) nie udało się przeprowadzić badania z żadnym uczniem.

Jako szkołę zastępczą należy wybrać wolną szkołę o pozycji szkoły w bazie najbliższej pozycji w bazie szkoły oryginalnie wylosowanej do badania.

Jeśli są dwie wolne szkoły o równie bliskiej pozycji szkoły w bazie (a więc równie odległych w bazie - jednak powyżej szkoły oryginalnej, druga poniżej) - należy wybrać tę o niższej pozycji w bazie.

"Wolna szkoła" to szkoła, która nie została wylosowana do badania, nie odmówiła udziału w badaniu, ani nie została dotąd wskazana jako szkoła zastępcza w innym przypadku
Uwaga: zawsze porównujemy pozycję w bazie możliwej szkoły zastępczej z pozycją w bazie szkoły pierwotnie wskazanej do badania (a nie z poprzednią szkołą zastępczą, w której nie udało się przeprowadzić badania).

3.3.5. Technika losowania szkół

- Tworzymy – na podstawie danych SIO – listę szkół (osobno w warstwach explicite)
- Dla każdej szkoły wyznaczamy dominujący obszar kształcenia (odpowiadający modalnej obszarów kształcenia uczniów danej szkoły)
- Dla każdej szkoły wyznaczamy szacowaną liczbę uczniów kwalifikujących się do badania , czyli uczniów nie starszych niż 29 lat w ostatniej klasie/ostatnim roku

nauki. Następnie eliminujemy z listy szkoły o szacowanej liczbie uczniów nie starszych niż 29 lat w ostatniej klasie/ostatnim roku nauki mniejszej niż 10.

- Dla każdej szkoły wyznaczamy wagę losowania równą
 - 1 dla szkół mających mniej niż 100 uczniów nie starszych niż 29 lat w ostatniej klasie/ostatnim roku nauki
 - $(100/[\text{liczba uczniów kwalifikujących się do badania}])$

oraz przypisujemy szkole wagę korygującą równą odwrotności wagi losowania.

- Porządkujemy hierarchicznie listę szkół ze względu na:
 - Dominujący obszar kształcenia
 - Województwo
 - Klasę wielkości miejscowości
 - Liczbę uczniów kwalifikujących się do badania

- Przypisujemy szkołom pozycję na uporządkowanej liście. Następnie rozdzielamy listę na dwie części: szkoły z numerami nieparzystymi wyłączamy z listy (będziemy z niej losowali szkoły do badania CAPI), a dalsze procedury prowadzimy jedynie na szkołach o numerach parzystych.

- Wyznaczamy krok losowania, będący ilorzem sumy wag losowania wszystkich szkół oraz liczby szkół przewidzianych do wylosowania w warstwie.

- Z użyciem generatora liczb pseudolosowych (rozkładu równomiernego) losujemy punkt startowy: iloczyn liczby losowej z przedziału $[0, 1]$ i [kroku losowania].

- Kolejnym szkołom przypisujemy skumulowaną wagę losowania:
 - Pierwszej szkole na liście przypisujemy wartość odpowiadającą równą jej wadze losowania powiększonej o wartość punktu startowego
 - Każdej kolejnej n-tej ($n > 1$) szkole na liście przypisujemy skumulowaną wagę losowania równą skumulowanej wadze losowania szkoły n-1 powiększonej o wagę losowania szkoły n.

- Wylosowane do próby są szkoły, dla których spełniony jest warunek:

$\text{Zaokrąglenie.w.dół}(SWL/[\text{krok}]) - \text{Zaokrąglenie.w.dół}(SWL-1/[\text{krok}]) > 0$

gdzie: SWL – skumulowana waga losowania danej szkoły; SWL-1 – skumulowana waga losowania poprzedniej szkoły na liście, [krok] – krok losowania.

Procedurę tę prowadzimy osobno we wszystkich trzech warstwach explicite.

3.3.6. Dobór uczniów do badania

Do badania audytoryjnego przewidziani są wszyscy uczniowie nie starsi niż 29 lat w ostatniej klasie/ostatnim roku nauki, nie ma zatem potrzeby formułowania dodatkowej reguły wyboru.

3.4. Koncepcja doboru próby: badanie CAPI

3.4.1. Warstwy explicite

Analogicznie jak w badaniu audytoryjnym, przyjmując jedynie inne liczebności szkół w warstwach:

- Warstwa średnich szkół zawodowych (techników): 300 szkół,
- Warstwa zasadniczych szkół zawodowych: 150 szkół,
- Warstwa szkół policealnych: 150 szkół.

Tym samym, warstwa techników, będąca relatywnie mniej intensywnie badana w badaniu audytoryjnym, będzie warstwą badaną najintensywniej w badaniu CAPI.

3.4.2. Warstwy implicite

Identyczne jak w badaniu audytoryjnym.

3.4.3. Wagi (prawdopodobieństwa wylosowania) poszczególnych szkół

Dla badania CAPI przyjęto zasadę stałej liczby uczniów do przebadania. Dla uzyskania odpowiedniej wielkości próby uczniów (co najmniej 15.000 uczniów wylosowanych), liczbę uczniów do przebadania w jednej szkole proponuję ustalić na 30 uczniów (prowadzi to do uzyskania próby uczniów liczącej około 16.400 uczniów). Aby uzyskać próbę, dla której każdy uczeń będzie miał jednakowe prawdopodobieństwo wylosowania, waga losowania danej szkoły powinna:

- być równa 30, jeśli liczba uczniów kwalifikujących się do badania jest w danej szkole mniejsza niż 30;
- być równa liczbie uczniów kwalifikujących się do badania w danej szkole w pozostałych przypadkach.

Podobnie jak w badaniu audytoryjnym, zaproponowano wprowadzenie dwóch modyfikacji motywowanych względami kosztowo-logistycznymi:

- Wyłączenia z badania szkół mających mniej niż 10 uczniów kwalifikujących się do badania.
- Przyjęcia jako maksimum wagi losowania wartości 100. Oprócz względów praktycznych, przywołanych przy omawianiu próby dla badania audytoryjnego, w przypadku badania CAPI potrzeba taka wynika ze zbyt małej liczby istniejących szkół liczących ponad 100 uczniów (zwłaszcza, gdy wziąć pod uwagę potrzebę zapewnienia szkół rezerwowych).
- Aby zmniejszenie prawdopodobieństwa ich doboru nie prowadziło do obciążenia uzyskanych w badaniu wyników, uczniowie pochodzący z tych szkół uzyskają w analizie danych dodatkową wagę korygującą równą, analogicznie jak w badaniu audytoryjnym, ($\frac{\text{liczba uczniów kwalifikujących się do badania}}{100}$).

3.4.4. Dobór szkół rezerwowych

Dobór szkół rezerwowych identyczny jak w przypadku badania audytoryjnego, z wykorzystaniem tej samej listy szkół rezerwowych.

3.4.5. Technika losowania szkół

Analogiczna jak dla badania audytoryjnego, z następującymi różnicami:

- Losowanie w oparciu o listę szkół o nieparzystych pozycjach w bazie.
- Wagi losowania :
 - 30 dla szkół mających mniej niż 30 uczniów kwalifikujących się do badania (uczniów nie starszych niż 29 lat w ostatniej klasie/ostatnim roku nauki).
 - Liczba uczniów kwalifikujących się do badania, tam, gdzie wynosi ona od 30 do 100.
 - 100, gdy $[\text{liczba uczniów kwalifikujących się do badania}] > 100$.

oraz przypisujemy szkole wagę korygującą równą:

- $\frac{[\text{liczba uczniów kwalifikujących się do badania}]}{100}$, jeśli $[\text{liczba uczniów kwalifikujących się do badania}] > 100$.
- 1 w pozostałych przypadkach.

3.4.6. Dobór uczniów do badania

Jeśli w szkole wskazanej do badania jest nie więcej, niż 30 uczniów kwalifikujących się do badania, należy przebadać wszystkich.

Jeśli jest ich więcej niż 30, należy przeprowadzić dobór techniką losowania systematycznego, z wykorzystaniem warstwowania implicite ze względu na: zawód kształcenia i wiek ucznia. Ze względu na możliwe trudności w uzyskaniu listy

zawierającej szczegółowe dane o zawodzie kształcenia, dopuszczalne jest oparcie losowania uczniów o danych o przypisaniu uczniów do oddziałów (które w większości przypadków – za wyjątkiem jedynie „klas wielozawodowych” – będą niosły tę samą informację).

Procedura winna przebiegać następująco:

- Należy uzyskać ze szkoły listę wszystkich uczniów ostatniego roku/ostatniej klasy urodzonych w latach 1988-2001.
- Lista powinna zawierać:
 - imię i nazwisko, ewentualnie inny jednoznaczny identyfikator (np. nazwę oddziału i numer w dzienniku);
 - rok i miesiąc urodzenia;
 - nazwę oddziału lub nazwę zawodu kształcenia.
- Listę uczniów należy uporządkować (hierarchicznie) ze względu na:
 - nazwę oddziału lub nazwę zawodu kształcenia (alfabetycznie);
 - rok i miesiąc urodzenia;
 - identyfikator ucznia (alfabetycznie).

Wyznaczamy krok losowania (będący ilorazem liczby uczniów na liście oraz liczby 30). Wyznaczamy punkt startowy będący iloczynem kroku losowania i liczby losowej z przedziału [0,1]. Dalej postępujemy analogicznie jak przy losowaniu szkół (przyjmując wagę losowania każdego ucznia równą 1).

Losowanie to (wyznaczania numerów na liście uczniów wylosowanych do próby) można również przeprowadzić za pomocą załączonego excelowego "Narzędzia losowania do CAPI".

W tym celu:

- w polu "liczba uczniów na liście" wpisujemy liczbę wszystkich uczniów na liście uczniów;
- w polu "pozycja szkoły w bazie" wpisujemy pozycję szkoły, dla której przeprowadzamy losowanie;
- do badania wylosowani są uczniowie, zajmujący na liście uczniów pozycje podane w komórkach poniżej.

3.5. Dobór próby a ważenie danych

3.5.1. Ważenie poststratyfikacyjne – wagi korygujące

W systemie wag należy zawsze uwzględniać czynnik *wagi korygującej* w_k wyznaczonej dla każdej dużej szkoły (powyżej 100 uczniów). Uwaga: w przypadku zastępowania szkół, zawsze jako wagę korygującą bierzemy wagę szkoły *oryginalnie wylosowanej do próby* (a więc jeśli oryginalna szkoła liczy ponad 100 uczniów kwalifikujących się do badania, ale została zastąpiona szkołą mniejszą – wagę korygującą uwzględniamy, natomiast jeśli oryginalna szkoła liczy mniej niż 100 uczniów kwalifikujących się do badania i została zastąpiona szkołą większą – wagi korygującej nie uwzględniamy).

3.5.2. Ważenie postrealizacyjne - wagi ze względu na poziom realizacji próby

Dla uczniów każdej szkoły należy wyznaczyć wagę wynikającą z poziomu realizacji w_r , równą ilorazowi liczby uczniów przewidzianych do badania oraz liczby uczniów uczestniczących w badaniu.

3.5.3. Ważenie wieńcowe

Ponieważ reprezentatywność ze względu na obszar kształcenia oraz ze względu na województwo było szczególnym postulatem zamawiającego, proponuję uwzględnienie tych zmiennych na kolejnym etapie ważenia, w procedurze ważenia wieńcowego. Ważenie wieńcowe powinno być przeprowadzone ze względu na rozkłady brzegowe województwa oraz obszaru kształcenia i powinno dotyczyć zbioru zważonego już wagami korygujących, oraz po korekcie wag ze względu na poziom realizacji. Ważenie wieńcowe powinno być przeprowadzone osobno w warstwach explicite. Ze względu na niewielką liczebność uczniowie niektórych obszarów kształcenia na poszczególnych typów szkół, na użytek ważenia wieńcowego proponuję połączyć:

- dla szkół policealnych: obszary budowlany i mechaniczno-górnicy z elektroniczno-elektryczny, oraz artystyczny z administracyjno-usługowym;
- dla techników: medyczno-społeczny i artystyczny z administracyjno-usługowym;
- dla ZSZ: rolniczo-leśny z mechanicznym oraz artystyczny z administracyjno-usługowym².

3.5.4. Typy szkół

Dobór próby dokonywany był przy założeniu odrębnej analizy danych w poszczególnych warstwach explicite, a więc osobnej analizy dla uczniów techników, szkół zasadniczych oraz szkół policealnych. Zwłaszcza w przypadku szkół policealnych można mieć bowiem

² Rozkłady brzegowe do ważenia wieńcowego wynikające z powyższych połączeń i danych Systemu Informacji Oświatowej umieszczone są w Tabelach 6.1. i 6.2. w rozdziale opisującym konstrukcję wag.

wątpliwości, co do ich rzeczywistej liczby: liczba uczniów wynikająca z danych SIO może znacząco odbiegać od liczby uczniów realnie uczestniczących w zajęciach i rzeczywiście dążących do zdobycia dzięki nim kwalifikacji zawodowych. Jednocześnie, trudno wskazać jakiegokolwiek alternatywne względem SIO źródło informacji o liczbie uczniów szkół policealnych.

W tej sytuacji zalecać muszę maksymalną ostrożność przy próbach łącznej analizy danych w zbiorze agregującym wyniki dla różnych typów szkół.

Na podstawie danych SIO, można oszacować liczby uczniów w wieku do 29 lat ostatnich klas poszczególnych typów szkół, jako:

Tabela 3.1. Rozkład liczebności i częstości uczniów ostatnich klas w wieku do 29 lat w poszczególnych typach szkół

Typ szkoły	Liczba uczniów do 29 lat	Częstość
Szkoła policealna (ponadgimnazjalna)	70 257	$P_p=29,85\%$
Technikum	116 124	$P_T=49,33\%$
Zasadnicza szkoła zawodowa	49 022	$P_z=20,82\%$

Źródło: obliczenia własne na podstawie danych Systemu informacji Oświatowej (wg stanu na dzień 30.09.2017 r.).

W przypadku analizowania łącznego zbioru danych system wag musi prowadzić do uwzględnienia powyższych proporcji, poprzez uwzględnienie w systemie wag czynnika będącego ilorzem odsetka uczniów danego typu szkół wg powyższej tabeli oraz ilorazu sumy wag dla uczniów danego typu szkół przez sumę wag dla wszystkich elementów próby.

3.6. Szacowanie błędów standardowych z uwzględnieniem schematu doboru próby

Ze względu na dostępność w zbiorze zmiennych kodujących informację o:

- warstwie explicite
- przypisaniu ucznia do wiązki realizacyjnej, jaką stanowi wylosowana szkoła

do szacowania błędów standardowych proponuję zastosowanie metod opartych na linearyzacji Taylora, odwołujących się do informacji o warstwie i wiązce. Metody te dostępne są w stosowanych pakietach do analizy danych statystycznych (STATA, AMStat).

3.7. Operat losowania

Operat losowania został skonstruowany według przyjętej definicji populacji na podstawie danych pobranych z Systemu Informacji Oświatowej według stanu na dzień 30.09.2017.

4. Narzędzia badawcze

Konstrukcja narzędzi badawczych poprzedzona była zapoznaniem się z założeniami i narzędziami stosowanymi w czterech wielkoskalowych międzynarodowych badaniach losów absolwentów oraz 14 krajowych programów badań losów absolwentów szkół zawodowych realizowanych w 11 europejskich i trzech pozaeuropejskich krajach. Na podstawie tej analizy³ stworzono narzędzia badawcze użyte w modułach sondażowych projektu MLEZAIMD. W badaniach tych wykorzystano cztery narzędzia badawcze. Były to:

- Kwestionariusz do wywiadu z dyrektorem szkoły,
- Ankieta audytoryjna dla ucznia,
- Kwestionariusz do wywiadu z uczniem,
- Kwestionariusz do wywiadu z absolwentem.

Kwestionariusze przekazano beneficjentowi projektu konkursowego, których w przypadku kwestionariuszy do wywiadów stworzył skrypty w oprogramowaniu CADAS, służącym do komputerowego wspomaganie wywiadu przeprowadzanego przez ankietera (CAPI – *Computer Assisted Personal Interview*). Skrypty i ich odpowiedniość względem przekazanych wersji kwestionariuszy (wyświetlanie poszczególnych treści, instrukcje dla ankietera, reguły przejścia, pętle itp.) zostały sprawdzone przez zespół badawczy IBE. Poszczególne wizualne elementy ankiety audytoryjnej zostały przygotowane w taki sposób, by uczniowie możliwie najlepiej rozumieli intencje badawczy, np. dotyczące możliwości zaznaczania jednej lub kilku odpowiedzi w danym pytaniu i umieli zastosować reguły przejścia.

Oprócz kwestionariuszy przygotowano również formularze, za pomocą których zbierano dane przed przystąpieniem uczniów do badania. Formularz wypełniał uczeń lub, jeżeli uczeń był niepełnoletni, jeden z jego rodziców lub opiekun prawny. Formularz zastosowano zarówno w przypadku badania audytoryjnego jak również modułu wywiadów z uczniami. Pierwsza część formularza zawierała pola, które wypełniane były podstawowymi danymi identyfikującymi respondenta (oraz rodzica/opiekuna prawnego w przypadku uczniów niepełnoletnich), łącznie z jego adresem domowym, numerem telefonu i adresem e-mail. Formularz zawierał do zaznaczenia pole, w którym wyrażano zgodę na udział ucznia w badaniu. W przypadku modułu wywiadów z uczniami zgoda dotyczyła również przetwarzania danych w celu ponownego kontaktu z respondentem. Kolejne pole dotyczyło zgody na pozyskanie przez Instytut Badań Edukacyjnych danych dotyczących aktywności edukacyjnej i zawodowej gromadzonych w bazach danych

³ Szczegółowa analiza zawarta jest w dwóch ekspertyzach przygotowanych na potrzeby projektu: Koniewski M., *Ekspertyza dotycząca metodologii badań sondażowych wykorzystywanych w badaniach losów absolwentów* oraz Koniewski M., *Ekspertyza dotycząca wskaźników i metod analizy statystycznej wykorzystywanych w badaniach losów absolwentów*.

instytucji publicznych. Wszystkie wersje formularzy zgody na udział w badaniu oraz przetwarzanie danych umieszczone są w aneksie do niniejszego raportu.

4.1. Kwestionariusz wywiadu z dyrektorem szkoły

Narzędzie do wywiadu z dyrektorem szkoły było przeznaczone dla wszystkich dyrektorów wylosowanych szkół, niezależnie od tego czy szkoła znalazła się w module badania audytoryjnego uczniów czy w module wywiadów z uczniami. W obu przypadkach kwestionariusz do wywiadu wyglądał tak samo. Kwestionariusz, który był podstawą do stworzenia skryptu komputerowego używanego w trakcie przeprowadzania wywiadu, zamieszczony jest w Aneksie 2a. W kwestionariuszu tym poruszone zostały następujące kwestie:

- podstawowe dane dotyczące respondenta: płeć, liczba lat pracy w danej szkole, liczba lat pracy jako dyrektor danej szkoły,
- podstawowe informacje o szkole: przynależność do zespołu, typy szkół wchodzące w skład zespołu, placówki kształcenia praktycznego wchodzące w skład zespołu,
- opis zawodów, w jakich prowadzone jest kształcenie w szkole, w tym: kod zawodu, liczba uczniów uczących się w danym zawodzie, kształcenie młodocianych pracowników i ich liczba, miejsce odbywania zajęć praktycznych oraz praktyk zawodowych,
- zawody, w stosunku do których zrezygnowano z rekrutacji lub rozpoczęto rekrutację w ciągu ostatnich trzech lat,
- projekty, w których szkoła brała udział w ciągu ostatnich trzech lat, oraz działania, zakupy, szkolenia realizowane w ramach tychże projektów,
- ocena infrastruktury technicznej szkoły, fakt posiadania pracowni lub warsztatów służących praktycznej nauce zawodu lub odbywaniu praktyk oraz ocena ich wyposażenia,
- możliwość pokrycia kosztów zakupów, remontów i innych działań za pomocą środków finansowych, którymi szkołą dysponuje,
- szczegółowe dane liczbowe dotyczące kadry nauczycielskiej przedmiotów ogólnych, teoretycznej i praktycznej nauki zawodu, w tym fakt zatrudnienia w oparciu o Kartę Nauczyciela, zatrudnienie na pełen etat, zatrudnienie w oparciu o inne formy zatrudnienia,
- fakt opracowywania i realizowania własnych programów nauczania w zawodzie,
- fakt występowania problemów z zatrudnianiem nauczycieli,
- miejsca odbywania przez uczniów zajęć praktycznych oraz praktyk zawodowych,

- fakt współpracy z pracodawcami, liczba pracodawców, z którymi szkoła współpracuje w obecnym roku szkolnym,
- w przypadku trzech pracodawców, z którymi współpracuje ma największe znaczenie: długość, regularność, charakter współpracy, wielkość firmy, fakt zatrudniania absolwentów, sformalizowanie współpracy,
- ocena zajęć praktycznych i praktyk zawodowych w przygotowaniu uczniów do wykonywania pracy na dobrym poziomie,
- ocena czynników, których poprawa w największym stopniu przyczyniłaby się do zwiększenia szans absolwentów szkoły na osiągnięcie sukcesu zawodowego,
- ocena szans absolwentów szkoły w porównaniu z innymi absolwentami danego typu szkoły w wyuczonym zawodzie, ocena ryzyka bezrobocia absolwentów szkoły,
- przewidywanie liczby miesięcy do znalezienia pierwszej pracy przez absolwentów oraz przeciętnego wynagrodzenia absolwenta w pierwszej pracy,
- ocena szans na kontynuowanie przez absolwentów nauki na studiach wyższych,
- źródła wiedzy o losach absolwentów na rynku pracy,
- dane dotyczące wysokości budżetu szkoły z podziałem na rok (2015, 2016, 2017), w tym budżet szkół w zespole kształcących zawodowo, dochody własne, środki publiczne,
- dane dotyczące liczby oddziałów, uczniów, personelu pedagogicznego oraz wszystkich zatrudnionych w całej szkole.

4.2. Ankieta audytoryjna

Narzędzie do sesji audytoryjnych z uczniami dystrybuowane było w formie papierowej. Każdy z uczniów otrzymywał cztery strony. Na pierwszej z nich przedstawione były podstawowe informacje o badaniu i prośba o wzięcie w nim udziału. Kolejne trzy zawierały pytania do wypełnienia przez uczniów. Narzędzie zawierało specjalne znaczniki, dzięki którym w późniejszym terminie beneficjent projektu konkursowego po zeskanowaniu ankiet, dzięki automatycznego procesowi ich sczytywania, stworzył część zbioru danych. Umożliwiło to także szybsze zakodowanie odpowiedzi otwartych oraz ustalenie faktycznie zaznaczonych przez respondentów odpowiedzi, które po sczytaniu przez program były niejednoznaczne. Ankieta audytoryjna stanowi Aneks 2b do niniejszego raportu. Zostały w niej poruszone następujące kwestie:

- podstawowe dane o respondencie: płeć, dokładna data urodzenia, zawód, którym respondent się kształci,

- fakt odbywania zajęć praktycznych oraz praktyk zawodowych: u pracodawcy, w szkole i w CKU/CKP,
- fakt zdania egzaminów (wszystkich lub części) potwierdzających kwalifikacje zawodowe niezbędne do uzyskania dyplomu,
- w przypadku nie zdania lub zdania części egzaminów: zamiary dotyczące przystąpienia do egzaminów potwierdzających kwalifikacje zawodowe,
- zamiar kontynuacji nauki w innej szkole lub na uczelni po ukończeniu nauki w obecnej szkole,
- zamiar podjęcia lub kontynuowania pracy w zawodzie, w którym respondent się kształci,
- najniższa miesięczna pensja, jaką respondent by zaakceptował w swojej pierwszej pracy po ukończeniu obecnej szkoły,
- spodziewana miesięczna pensja za dwa lata,
- planowanie wyjazdu zagranicznego związanego z nauką lub pracą w ciągu najbliższych 12 miesięcy,
- liczba prac dorywczych lub sezonowych podjętych od ukończenia gimnazjum,
- łączna liczba miesięcy zatrudnienia w pracach trwających dłużej niż jeden miesiąc od ukończenia gimnazjum do chwili obecnej,
- fakt ograniczenia funkcjonowania spowodowanego poważną chorobą lub niepełnosprawnością,
- fakt powtarzania klasy kiedykolwiek w trakcie nauki w obecnej szkole lub wcześniej,
- wykształcenie rodziców lub opiekunów prawnych.

Dodatkowo dane identyfikacyjne na ankiecie służą określeniu typu szkoły, do której uczeń uczęszcza.

4.3. Kwestionariusz do wywiadu z uczniem

Narzędzie do wywiadu z uczniem było używane w szkołach wylosowanych do próby w module CAPI. W każdej takiej szkole było ono przeznaczone dla 30 uczniów wylosowanych spośród wszystkich uczniów ostatniej klasy mających nie więcej niż 29 lat. Kwestionariusz, który był podstawą do stworzenia skryptu komputerowego używanego w trakcie przeprowadzania wywiadu, zamieszczony jest w Aneksie 2c. W kwestionariuszu tym poruszone zostały następujące kwestie:

- podstawowe dane o respondencie: płeć, dokładna data urodzenia, typ szkoły, typ oddziału – wielozawodowy, zawód, w którym respondent się kształci,
- w przypadku uczniów szkół policealnych fakt zdania matury,
- blok dotyczący historii edukacyjnej: doradztwo zawodowe w gimnazjum, samoocena w ostatniej klasie gimnazjum (dla szkół policealnych: w szkole ponadgimnazjalnej), sposób trafienia do obecnej szkoły, kryteria wyboru szkoły, kryteria wyboru zawodu, w którym uczeń się kształci, ocena obecnej szkoły, fakt odbywania zajęć praktycznych oraz praktyk zawodowych: u pracodawcy, w szkole i w CKU/CKP, hipotetyczny ponowny wybór tej samej szkoły i tego samego zawodu,
- blok dotyczący kursów, szkoleń, zdawania dodatkowych egzaminów lub certyfikatów od czasu ukończenia gimnazjum: typ szkoleń, instytucja prowadząca, fakt ukończenia szkolenia, data ukończenia, uzyskanie dyplomu lub certyfikatu, fakt odpłatności za szkolenie lub egzamin,
- blok praca: liczba prac dorywczych lub sezonowych podjętych od ukończenia gimnazjum, liczba tygodni poświęconych na takie prace, fakt otrzymywania formalnych umów, obszar podejmowania takich prac, dla każdej z prac trwających dłużej niż jeden miesiąc od ukończenia gimnazjum do chwili obecnej: miejsce pracy, okres zatrudnienia, forma zatrudnienia, lokalizacja względem domu, fakt otrzymywania wynagrodzenia większego niż zapisane w umowie,
- blok dotyczący planów na przyszłość: zamiar przystąpienia do egzaminu maturalnego, zamiar do przystąpienia do brakujących egzaminów zawodowych (jeśli takie są), powody odstąpienia od takiego zamiaru, zamiar kontynuacji nauki w innej szkole lub na uczelni po ukończeniu nauki w obecnej szkole (dane szkoły, uczelni, kierunku, zawodu), zamiar podjęcia lub kontynuowania pracy w zawodzie, w którym respondent się kształci, zamiar poszukiwania pracy w ciągu 3 miesięcy po ukończeniu szkoły, charakter poszukiwanej pracy, poszukiwanie pracy zagranicą, oczekiwania wobec pierwszej pracy, przewidywania daty podjęcia pierwszej pracy po szkole, przewidywanie czy będzie to praca w wyuczonym zawodzie, najniższa miesięczna pensja, jaką respondent by zaakceptował w swojej pierwszej pracy po ukończeniu obecnej szkoły, spodziewana miesięczna pensja za dwa lata,
- blok dotyczący charakterystyki respondenta: typ miejscowości zamieszkania, zaspokajanie potrzeb finansowych, regularne wsparcie finansowe od rodziny, stan cywilny, fakt posiadania dzieci, liczba dzieci, liczba osób w gospodarstwie domowym, dla każdej z osób: pokrewieństwo, rok urodzenia, płeć, fakt uczenia się lub studiowania oraz podejmowania pracy; sposób gospodarowania pieniędzmi, fakt posiadania gospodarstwa rolnego przez kogoś z gospodarstwa domowego (powierzchnia gruntów), liczba rodzeństwa, planowanie wyjazdu zagranicznego związanego z nauką lub pracą w ciągu najbliższych 12 miesięcy, fakt podejmowania pracy zagranicą przez kogoś z rodziny lub najbliższych znajomych, liczba książek w domu rodzinnym w wieku 10 lat, miesięczna kwota do dyspozycji respondenta, fakt powtarzania klasy kiedykolwiek w trakcie nauki

w obecnej szkole lub wcześniej, wykształcenie rodziców lub opiekunów prawnych, ocena stanu zdrowia, fakt ograniczenia funkcjonowania spowodowanego poważną chorobą lub niepełnosprawnością.

Po rozmowie ankietier potwierdzał wyświetloną przez skrypt datę i godzinę zakończenia wywiadu oraz jego czas lub korygował dane komentując dodatkowe okoliczności, przez które dane nie odpowiadają rzeczywistości. Następnie zaznaczał, czy respondent miał jakieś kłopoty ze zrozumieniem pytań, czy miał problemy z operowaniem materiałem liczbowym (kwotami, datami) oraz zaznaczał swoją płeć. Następnie zgodnie z założeniem beneficjenta projektu konkursowego wprowadzał również dane z formularza zgody na udział w badaniu respondenta lub rodzica, by każdy wywiad był przyporządkowany do odpowiedniego ucznia.

4.4. Kwestionariusz do wywiadu z absolwentem

Narzędzie do wywiadu z absolwentem było używane w trakcie wizyty w domu respondenta. Próbę stanowili absolwenci szkół, z którymi został przeprowadzony wywiad rok wcześniej, gdy byli jeszcze uczniami. Kwestionariusz, który był podstawą do stworzenia skryptu komputerowego używanego w trakcie przeprowadzania wywiadu, zamieszczony jest w Aneksie 2d. W kwestionariuszu tym poruszone zostały następujące kwestie:

- potwierdzenie danych respondenta, kontrola faktu uczęszczania do konkretnej szkoły, kształcenia się w konkretnym zawodzie,
- fakt ukończenia szkoły, data uzyskania świadectwa lub matury, powody nieukończenia, termin przerywania nauki
- fakt uzyskania dyplomu potwierdzającego wszystkie kwalifikacje w zawodzie, fakt podchodzenia do egzaminów potwierdzających zdobycie kwalifikacji, powody nie przystąpienia do egzaminów, powody nie uzyskania kwalifikacji, zamiar przystąpienia do egzaminów,
- fakt podjęcia nauki w liceum dla dorosłych: data rozpoczęcia, ukończenia lub przerywania nauki, odpłatność, zamiar przystąpienia do matury, tryb nauki i odpłatność,
- fakt składania podania na studia wyższe: liczba kierunków, nazwy kierunków i uczelni, fakt przyjęcia, fakt rozpoczęcia, ukończenia lub przerywania nauki, uzyskanie dyplomu, tryb nauki i odpłatność, powody podjęcia nauki na studiach,
- fakt składania podania o przyjęcie do szkoły policealnej: liczba szkół, nazwy kierunków, zawodów i miejscowości, w której znajduje się szkoła, fakt przyjęcia, fakt rozpoczęcia, ukończenia lub przerywania nauki, uzyskanie dyplomu, tryb nauki i odpłatność, powody podjęcia nauki w szkole policealnej,

- blok dotyczący kursów, szkoleń, zdawania dodatkowych egzaminów lub certyfikatów w trakcie nauki szkolnej lub po jej ukończeniu: typ szkoleń, instytucja prowadząca, fakt ukończenia szkolenia, data ukończenia, uzyskanie dyplomu lub certyfikatu, fakt odpłatności za szkolenie lub egzamin,
- blok praca: liczba prac dorywczych lub sezonowych podjętych od ukończenia szkoły (lipiec 2017), liczba tygodni poświęconych na takie prace, fakt otrzymywania formalnych umów, obszar podejmowania takich prac, dla każdej z prac trwających dłużej niż jeden miesiąc od ukończenia gimnazjum do chwili obecnej: miejsce pracy, okres zatrudnienia, forma zatrudnienia, lokalizacja względem domu, czas dojazdu z domu do pracy, fakt otrzymywania wynagrodzenia większego niż zapisane w umowie,
- blok pytań o pierwszą i ostatnią pracę: długość okresu poszukiwania pracy po ukończeniu szkoły, zgodność pracy z wyuczonym zawodem, nazwa stanowiska i opis obowiązków, wymiar czasu pracy i liczba godzin w ciągu tygodnia w tym miejscu pracy, wysokość miesięcznych zarobków, wielkość firmy/instytucji zatrudniającej respondenta, ewentualnie: powód zakończenia pracy w danym miejscu,
- zadowolenie z obecnej pracy w kontekście oczekiwań pod względem różnych czynników, ocena łatwości znalezienia podobnej lub lepszej pracy u innego pracodawcy, ocena łatwości znalezienia przez pracodawcę pracownika na miejsce respondenta, obawa przed utratą pracy, poszukiwanie innej lub dodatkowej pracy w ciągu ostatnich 4 tygodni, powody poszukiwania pracy,
- blok bezrobocie: fakt bycia bez pracy, okres bycia bezrobotnych, fakt poszukiwania pracy, fakt bycia zarejestrowanym jako bezrobotny w urzędzie pracy, powód ostatniej rejestracji w urzędzie pracy, fakt korzystania z pomocy urzędu w różnych formach, otrzymanie oferty pracy z urzędu, podjęcie tej pracy, powody niepodjęcia tej pracy,
- przewidywanie dotyczące sytuacji zawodowej respondenta za pięć lat w kontekście możliwości finansowych,
- blok dotyczący charakterystyki respondenta: płeć, rok urodzenia, typ miejscowości zamieszkania, województwo, powiat, gmina zamieszkania, fakt mieszkania w tej samej miejscowości, w której respondent się wychowywał, ewentualnie oddalenie od tej miejscowości, zaspokajanie potrzeb finansowych, regularne wsparcie finansowe od rodziny, stan cywilny, fakt posiadania dzieci, liczba dzieci, liczba osób w gospodarstwie domowym, dla każdej z osób: pokrewieństwo, rok urodzenia, płeć, fakt uczenia się lub studiowania oraz podejmowania pracy; sposób gospodarowania pieniędzmi, fakt posiadania gospodarstwa rolnego przez kogoś z gospodarstwa domowego (powierzchnia gruntów),
- dodatkowy blok dotyczący staży lub praktyk zawodowych w trakcie nauki szkolnej realizowanych zagranicą.

W przypadku gdy respondent z jakichś powodów nie brał udziału w wywiadzie ankieter miał za zadanie zaznaczyć jeden spośród 20 predefiniowanych powodów niezrealizowania wywiadu lub opisać sytuację bardziej szczegółowo.

5. Realizacja badania w terenie

Za realizację badania w terenie odpowiedzialna był beneficjent projektu wyłonionego przez Ministerstwo Edukacji Narodowej w drodze postępowania konkursowego. Była to firma PBS Sp. z o.o. Zadaniem firmy było przeprowadzenie badań terenowych na podstawie narzędzi i założeń badawczych skonstruowanych przez zespół badawczy Instytutu Badań Edukacyjnych.

5.1. Organizacja realizacji badania

Pierwszym etapem badania była realizacja dwóch modułów ilościowych w szkołach. Obydwa moduły wymagały podobnego charakteru działań organizacyjnych. W grudniu 2016 roku zespół badawczy Instytutu Badań Edukacyjnych przekazał narzędzia badawcze: kwestionariusz dyrektora szkoły, kwestionariusz ucznia oraz ankietę audytoryjną. Oprócz tego beneficjent projektu konkursowego otrzymał dostęp do narzędzia on-line, które zawierało próbę szkół wylosowanych do każdego z modułów, a także możliwość samodzielnego pobierania szkół zastępczych (w przypadku braku możliwości realizacji badania w wylosowanej szkole). Przekazano także narzędzie wykorzystywane w module wywiadów z uczniami, dzięki któremu, po wprowadzeniu danych wszystkich uczniów danej szkoły, którzy spełniają warunki do wzięcia udziału w badaniu, losowano założoną liczebność 30 uczniów.

W toku dalszych prac beneficjent projektu konkursowego stworzył skrypty w oprogramowaniu do realizacji wywiadów ze wspomaganie komputerowym (CAPI). Skrypty te podlegały kilkukrotnej weryfikacji dokonywanej przez zespół badawczy IBE. Wprowadzono do nich niezbędne korekty i ustalono szczegóły związane m.in. z kodowaniem poszczególnych zmiennych w zbiorach danych, które miały być efektem badań. Stworzono również ostateczny wygląd formularza ankiety papierowej, uwzględniając m.in. znaczniki niezbędne do założonej przez beneficjenta projektu konkursowego metody wprowadzania danych z ankiet za pomocą skanowania. Ustalono też inne, liczne szczegóły związane z realizacją obydwu modułów.

Rekrutacja szkół rozpoczynała się od wysłania listu zapowiadającego informującego dyrektora szkoły o zamiarze przeprowadzenia badania, jego celach i zakresie. Dalsza część rekrutacji prowadzona była telefonicznie i przy użyciu poczty elektronicznej. Scenariusz rekrutacji został ustalony z zespołem badawczym IBE i zawierał: wstępną zgodę dyrektora na udział szkoły w badaniu (ewentualnie wskazanie na powód odmowy), wskazanie osoby, która weźmie udział w wywiadzie opartym na kwestionariuszu dyrektora szkoły (procedura dopuszczała wskazanie przez dyrektora innego pracownika szkoły, który weźmie udział w wywiadzie, np. wicedyrektora), wyznaczenie koordynatora szkolnego do kontaktów w sprawie realizacji badania, możliwości infrastrukturalne przeprowadzenia sesji badawczych, wstępne preferencje dotyczące terminów przeprowadzenia badania, liczbę oddziałów ostatniej klasy (ostatniego roku kształcenia), zawody, w jakich odbywa się kształcenie w tych oddziałach oraz liczbę uczniów w każdym z oddziałów. W trakcie rozmowy dyrektor otrzymywał szczegółowe informacje na temat realizacji badania, które przesyłane były również za pomocą poczty elektronicznej. Po uzyskaniu pisemnej zgody dyrektora szkoły na przeprowadzenie

badania do danej szkoły przypisywany był jeden lub kilku ankieterów, którzy byli odpowiedzialni za realizację badania.

W ramach czynności przygotowawczych uruchomiono również infolinię dla szkół i uczniów, dzięki której można się było skontaktować z firmą odpowiedzialną za realizację i poznać szczegóły badania. Każdy z uczniów, który został wylosowany do badania (lub w przypadku ankiety audytoryjnej, który był w oddziale ostatniej klasy), otrzymał ulotkę informującą o celach, zakresie i przebiegu planowanego badania oraz formularz zgody na przeprowadzenie badania. W przypadku gdy uczeń był niepełnoletni otrzymywał on do przekazania rodzicom lub opiekunom prawnym specjalnie przygotowany formularz zgody rodzica/opiekuna prawnego na udział ucznia w badaniu wraz z listem intencyjnym informującym o celach, zakresie i przebiegu badania.

W materiałach informacyjnych uczniowie zostali poinformowani o upominkach przewidzianych dla uczestników badania. Dla szkół były to zestawy sportowe. Dla uczniów biorących udział w ankiecie audytoryjnej indywidualny dostęp do materiałów edukacyjnych z zakresu przedsiębiorczości, zarządzania karierą i rozwoju osobistego. Dla uczniów, którzy wezmą udział w obu etapach badania panelowego, przewidziano upominek w postaci audiobooka/e-booka do nauki języka obcego oraz możliwość wzięcia udziału w konkursie, w którym nagrodami będą rowery, tablety i czytniki książek.

Realizację terenową badania poprzedzało szkolenie ankieterów, które odbyło się 27 i 28 marca 2017 roku, na którym omówiono wszelkie niezbędne szczegóły realizacyjne związane m.in. z: rekrutacją szkół i uczniów, doбором respondentów, harmonogramem badań, dokumentami niezbędnymi do realizacji badania, formularzami zgód uczniów i rodziców, platformą do zarządzania badaniem, odsyłaniem materiałów itp. Główną częścią szkolenia było omówienie każdego narzędzia badawczego wraz z treścią pytań i omówieniem potencjalnych problemów w trakcie zbierania danych. Warunkiem pracy przy badaniu była obecność ankietera na szkoleniu.

Podobny schemat przygotowań nastąpił rok później i był związany z drugą falą badania podłużnego – badaniem absolwentów. Przygotowania obejmowały przygotowanie przez beneficjenta projektu konkursowego skryptu na podstawie tekstu narzędzia przekazanego przez zespół badawczy IBE, wysłanie listu intencyjnego i próbę umówienia się na wywiad z respondentami z pierwszej fali badania CAPI, którzy zgodzili się na powtórny kontakt, a także przeprowadzenie szkolenie dla ankieterów, biorących udział w badaniu absolwentów.

5.2. Techniki i techniki gromadzenia danych

Badanie terenowe w szkołach zaplanowano na drugi semestr roku szkolnego 2016/17. Ze względu na przygotowania organizacyjne i długi okres rekrutacji szkół do badania beneficjent projektu konkursowego rozpoczął realizację terenową 10 kwietnia 2017 roku z planem zakończenia badania wraz z końcem roku szkolnego.

Wywiady z dyrektorami w znacznej większości odbyły się między 10 kwietnia a 30 czerwca. Ze względu na problemy realizacyjne beneficjent projektu konkursowego przedłużył zbieranie danych na okres wakacyjny. 215 wywiadów z dyrektorami zostało

przeprowadzonych w lipcu i sierpniu. Średni czas trwania wywiadu z dyrektorem wyniósł 30 minut (mediana 27 minut).

Moduł audytoryjny badania odbył się w szkołach w kwietniu, maju i czerwcu. Przed badaniem zbierane były zgody uczniów (lub, w przypadku uczniów niepełnoletnich, zgody rodziców) na udział w badaniu. Wraz z formularzami zgody na udział w badaniu uczniowie otrzymywali broszury informacyjne zawierające cele, sposób przeprowadzenia badania oraz dane kontaktowe do wykonawcy badania. Ankieter przeprowadzał badanie we wcześniej ustalonym terminie w danym oddziale ostatniej klasy w trakcie jednej godziny lekcyjnej. W trakcie zwracania ankiet sprawdzał ich kompletność przypisywał poszczególne ankiety do wcześniej zebranych zgód używając do tego celu identyfikatorów nadanych każdemu z respondentów.

Wywiady z uczniami w module CAPI odbywały się na terenie szkoły i były przeprowadzane z uczniami, którzy zostali wylosowani do badania i dostarczyli zgody na udział w badaniu. W każdej szkole biorącej udział w tym module robiono spis uczniów należących do badanej populacji, a następnie spośród nich losowano 30 uczniów, których proszono o wzięcie udziału w badaniu i wypełnienie formularza zgody na udział w badaniu. Wraz z formularzami zgody na udział w badaniu uczniowie otrzymywali broszury informacyjne zawierające cele, sposób przeprowadzenia badania oraz dane kontaktowe do wykonawcy badania. Po przeprowadzonym wywiadzie ankieter wprowadzał do skryptu dane ze zgody ucznia. W przypadku tego modułu napotkano na największe problemy związane z realizacją założonej próby w trakcie roku szkolnego. Do końca czerwca zrealizowano jedynie 6949 wywiadów. W wakacje (w lipcu i sierpniu) zrealizowano dodatkowe 206 wywiadów. Po konsultacjach z instytucją zarządzającą projektem podjęto decyzję o przedłużeniu realizacji badania do grudnia 2017 i wykorzystaniu do ich realizacji kontaktów do uczniów zdobytych w trakcie badania audytoryjnego. Dodatkowe wywiady przeprowadzono od października do grudnia 2017 roku z 2947 uczniami. Średni czas trwania wywiadu z uczniem wyniósł 23 minuty (mediana 22 minut).

Badanie terenowe polegające na powrocie do absolwentów i przeprowadzeniu z nimi wywiadu odbywało się od 10 marca do 30 czerwca 2018 roku w domach respondentów lub w miejscach dla nich dogodnych po wcześniejszym umówieniu się z ankieterem. Do umówienia wywiadu wykorzystywano dane kontaktowe zebrane w trakcie pierwszego kontaktu z respondentem. Ze względu na opisane powyżej problemy z realizacją badania w szkołach oraz koniecznością upłynięcia większej ilości czasu pomiędzy dwoma wywiadami ustalono, że powrót do osób badanych pierwszym kwestionariuszem w końcówce 2017 roku nastąpi pod koniec realizacji badania, czyli w maju lub czerwcu 2018 roku. Średnia czas trwania wywiadu z absolwentem wynosił 21 minut, przy medianie wynoszącej również 21 minut.

Badanie wykorzystujące każde z narzędzi zostało poprzedzona pilotażem zrealizowanym na niewielkiej liczbie uczniów, absolwentów lub dyrektorów. Dzięki temu przetestowano ostatecznie skrypty wykorzystywane w realizacji badania oraz formularz ankiety. Ponadto sprawdzono czy długość przeprowadzania wywiadu zgadza się z założeniami, a także sprawdzono stopień zrozumienia treści pytań przez respondentów oraz sposób zapisu poszczególnych informacji w zbiorach danych.

5.3. Realizacja próby

Beneficjent projektu konkursowego realizował badanie w 1391 szkołach (w tym w: 506 ZSZ, 472 technikach i 396 szkołach policealnych). W 1374 z nich dyrektor (lub osoba przez niego wyznaczona) wziął udział w wywiadzie opartym na kwestionariuszu dyrektora szkoły (w tym w: 513 ZSZ, 476 technikach i 402 szkołach policealnych). W 828 szkołach co najmniej jeden uczeń wziął udział module wywiadów CAPI z uczniami (w tym w: 304 ZSZ, 345 technikach i 179 szkołach policealnych). W 679 szkołach przeprowadzono badanie metodą ankiety audytoryjnej (w tym w: 292 ZSZ, 143 technikach i 244 szkołach policealnych). Ze względu na opisane wyżej problemy z realizacją modułu wywiadów CAPI z uczniami zdarzało się, że uczniowie danej szkoły brali udział zarówno w module ankiety audytoryjnej, jak i w module wywiadów CAPI. Ogółem w jednym lub drugim module (lub w obu) wzięli udział uczniowie z 1216 szkół (w tym z: 467 ZSZ, 429 technikach i 320 szkół policealnych). Oznacza to, że w przypadku 175 szkół (46 ZSZ, 47 techników i 82 szkół policealnych) dysponujemy jedynie danymi jedynie z wywiadu z dyrektorem szkoły.

W module CAPI zbadano 10102 uczniów (ZSZ: 3848, technikum: 4972, szkoła policealna: 1282), a w module ankiety audytoryjnej 12 248 uczniów (ZSZ: 6071, technikum: 3549, szkoła policealna: 2628).

W badaniu absolwentów wzięło udział 7713 respondentów spośród 10102 uczniów biorących udział w badaniu w szkole (2834 wcześniejszych uczniów ZSZ, 3918 wcześniejszych uczniów techników oraz 961 wcześniejszych uczniów szkół policealnych). Poziom realizacji wywiadów z absolwentami stanowi ogółem 76,4% wcześniejszych wywiadów realizowanych w szkole. W podziale na typ szkoły response rate badania absolwentów wyniósł: dla ZSZ: 73,6%; dla techników: 78,8%; dla szkół policealnych 75,0%. Tabela 5.1. zawiera przyczyny braku realizacji wywiadu z absolwentem podane przez beneficjenta projektu konkursowego.

W związku z napotkanymi problemami w realizacji badania (przeprowadzenie badania w szkole, która nie została wylosowana, realizacja wywiadu z 1-2 uczniami w danej szkole spośród 30 wylosowanych) niektórym jednostkom obserwacji w zbiorach danych nadano zerową wagę. Szczegółowo opisano to w rozdziale 6.2.

Tabela 5.1. Przyczyny niezrealizowania wywiadu z absolwentem

	Liczebność	Odsetek spośród wszystkich kontaktów	Odsetek spośród wywiadów niezrealizowanych
Adres niewykorzystany – taki, pod który w ogóle się nie udano	632	6,3	26,5
Odmowa respondenta: rozmawiano z respondentem i odmówił on udziału w badaniu	555	5,5	23,2
Drzwi zamknięte	378	3,7	15,8
Pod wskazanym adresem nastąpił kontakt z osobą niebędącą respondentem, która stwierdziła, że respondent już tam nie mieszka i odmówiła udzielenia informacji o jego obecnym miejscu zamieszkania	229	2,3	9,6
Brak możliwości skontaktowania się osobiście z respondentem przed ukończeniem okresu realizacji badania (ze względu na dłuższy pobyt za granicą)	211	2,1	8,8
Pod wskazanym adresem nastąpił kontakt z osobą niebędącą respondentem, która stwierdziła, że respondent już tam nie mieszka, ale nie była w stanie udzielić informacji o jego obecnym miejscu zamieszkania	168	1,7	7,0
Wywiad przerwany (rozpoczęto realizację ze wskazaną osobą, ale nie ukończono wywiadu)	68	0,7	2,8
Wywiad usunięty w wyniku kontroli	62	0,6	2,6
Odmowa gospodarstwa: rozmawiano z kimś z gospodarstwa domowego, kto odmówił współpracy przy identyfikacji respondenta, i nie wiadomo, czy był to respondent	52	0,5	2,2
Brak możliwości skontaktowania się osobiście z respondentem przed ukończeniem okresu realizacji badania z innych powodów (np. przebywa w szpitalu, więzieniu, itp.)	16	0,2	0,7
Pod wskazanym adresem nastąpił kontakt z osobą niebędącą respondentem, która stwierdziła, że respondent już tam nie mieszka i udzieliła informacji o jego obecnym miejscu zamieszkania, numerze telefonu lub adresie e-mail	8	0,1	0,3
Brak informacji o respondencie (nie wiadomo, czy pod danym adresem ktoś mieszka)	3	0,0	0,1
Adres nie istnieje	2	0,0	0,1
Rejon niedostępny (niemożność dotarcia w okolice wywiadu, odludzie, niebezpieczna okolica, huragan)	1	0,0	0,0
Inne	4	0,0	0,2
Ogółem	10102	23,6	100,0

5.4. Problemy w badaniu terenowym

W trakcie realizacji badania beneficjent projektu konkursowego natrafił na liczne problemy w osiągnięciu wskaźników założonych we wniosku konkursowym. Dotyczyły one głównie:

- Kwestii organizacyjnych związanych z terminem rozpoczęcia badania, co wpłynęło szczególnie na bardzo krótki okres przeprowadzania badania w technikach, ze względu na nieodległy od terminu rozpoczęcia badania termin matur, do których podchodzili uczniowie należący do populacji badania.
- Niskiego poziomu realizacji próby na poziomie szkół, wynikającego z odmów dyrektorów szkół na udział w badaniu i konieczności stosowania prób rezerwowych (nawet kilku zamienników dla jednej szkoły wylosowanej do próby głównej). W szczególności dotyczyły one szkół policealnych, dla których skala odmów była tak duża, że wyczerpał się operat losowania. Oznacza to, że nastąpiła próba kontaktu ze wszystkimi szkołami policealnymi zawartymi w operacie stworzonych na podstawie Systemu Informacji Oświatowej.
- Problemów polegających na nieobecności uczniów w trakcie zajęć lekcyjnych, co skutkowało brakiem możliwości dostarczenia uczniowi formularza zgody na udział w badaniu i/lub brakiem możliwości wzięcia udziału w badaniu przez ucznia.
- Problemów z realizacją modułu wywiadów z uczniami, czego efektem było przeprowadzanie dodatkowej części badania pomiędzy październikiem i grudniem kolejnego roku szkolnego korzystając z kontaktów zebranych w trakcie przeprowadzania modułu audytoryjnego.

6. Przetwarzanie danych

Niniejszy rozdział zawiera opis procedur, które zostały wykonane w celu uzyskania zbiorów danych do analiz. Przedstawiona jest struktura zbiorów danych otrzymanych od beneficjenta projektu konkursowego, które zostały poddane dalszemu przetwarzaniu. Opis zawiera sposób wytworzenia wag wynikających z zastosowanego schematu doboru próby oraz efektów realizacyjnych. W dalszej części opisano dwie nowe struktury zbiorów danych wytworzone ze względu na formę pytań zadanych respondentom oraz idącą za tym konieczność specyficznego sposobu analizy części zagadnień umieszczonych w narzędziach badawczych:

- zbiór epizodów dotyczących historii edukacyjnej i zawodowej absolwentów,
- zbiory osobo-miesiący zawierający cechy respondentów w kolejnych momentach czasu, w przypadku tego badania w kolejnych miesiącach.

6.1. Struktura podstawowych zbiorów danych

Wszystkie zbiory danych z badania zostały wstępnie przygotowane przez beneficjenta projektu konkursowego przy użyciu poniższych zasad tworzenia zbiorów obowiązujących w Instytucie Badań Edukacyjnych:

1. Wraz z zanonimizowanymi zbiorach danych ilościowych zebranych w badaniu należy dostarczyć opisy ich struktury (codebook).
2. Zarówno zanonimizowane zbiory danych ilościowych, jak i opisy ich struktury powinny mieć format csv:
 - 2.1. separatorem pól powinien być średnik (kod ASCII 59);
 - 2.2. separatorem tekstu powinien być cudzysłów (kod ASCII 34);
 - 2.3. cudzysłów w tekście oznacza się przez następujące po sobie dwa znaki cudzysłów (kod ASCII 34) – w takim kontekście są one interpretowane jako pojedynczy cudzysłów, a nie dwa znaki separatora tekstu;
 - 2.4. znaki narodowe powinny być zakodowane w stronie kodowej CP-1250.
3. Opis struktury zanonimizowanego zbioru danych ilościowych (codebook) powinien zawierać:
 - 3.1. W pierwszym wierszu nagłówki kolumn (patrz punkt 3.3. poniżej).
 - 3.2. W kolejnych wierszach opis poszczególnych zmiennych zanonimizowanego zbioru danych ilościowych (patrz opis kolumn)

- 3.2.1. N-ty wiersz (licząc z pominięciem wiersza zawierającego nagłówki kolumn) musi opisywać N-tą zmienną zanonimizowanego zbioru danych ilościowych (a więc zmienną, której wartości znajdują się w N-tej kolumnie zanonimizowanego zbioru danych ilościowych).
- 3.3. W kolejnych kolumnach opis cech poszczególnych zmiennych zanonimizowanego zbioru danych ilościowych (dla każdej zmiennej muszą zostać podane wszystkie stosowne dla danego typu zmiennej informacje):
 - 3.3.1. nazwa zmiennej;
 - 3.3.1.1. musi być unikalna w obrębie zbioru danych;
 - 3.3.1.2. powinna oddawać hierarchię zmiennych, tzn. składać się z kodów sekcji/kategorii/grup zmiennych rozdzielonych znakiem podkreślenia. Np.: M_12_a – popunkt a pytania 12 w sekcji M ankiety;
 - 3.3.1.2.1. każdy kod sekcji/kategorii/grupy zmiennych, który zostanie użyty w nazwie jakiegokolwiek zmiennej musi posiadać w opisie struktury zanonimizowanego zbioru danych ilościowych zmienną typu kategoria, która opisuje daną sekcję/kategorię/grupę, np. jeśli zbiór zawiera zmienne S_1, S_2, M_1, M_2, to w opisie struktury zanonimizowanego zbioru danych ilościowych muszą pojawić się definicje zmiennych typu kategoria o nazwach S i M, których opis zawierać będzie nazwę danej sekcji/kategorii/grupy zmiennych, np. M – metryczka, S – opinie o sąsiadach (patrz też 3.3.5.1).
 - 3.3.1.3. dla danych pochodzących z ankiet kwestionariuszowych nazwa zmiennej powinna odpowiadać numerowi pytania w kwestionariuszu ankiety; w wypadku gdy jednemu pytaniu kwestionariusza odpowiada kilka zmiennych (np. pytanie z możliwością wielokrotnego wyboru), nazwy zmiennych odpowiadających danemu pytaniu powinny się składać z numeru pytania oraz sufiksu _N, gdzie N to kolejne liczby naturalne;
 - 3.3.1.4. nie może zawierać białych znaków (w szczególności spacji);
 - 3.3.2. krótki opis zmiennej, zwięźle (w maksymalnie 70 znakach) opisujący czego dana zmienna dotyczy;
 - 3.3.3. opis zmiennej, w jasny sposób definiujący, czego dana zmienna dotyczy;
 - 3.3.3.1. dla danych pochodzących z ankiet kwestionariuszowych powinna ona zawierać pytanie odczytywane respondentowi w ankiecie odpowiadające danej zmiennej;
 - 3.3.3.2. jeśli jest to zmienna wywiedziona z innych zmiennych, w tym polu musi zostać opisany sposób, w jaki została wywiedziona;
 - 3.3.4. lista słów kluczowych powiązanych ze znaczeniem zmiennej. Poszczególne wyrażenia oddzielone od siebie znakiem przecinka;

- 3.3.5. typ zmiennej: liczba całkowita/liczba rzeczywista/tekst/data/TERC / kategoria /waga ;
- 3.3.5.1. kategoria to specjalny typ, który opisuje jedynie strukturę narzędzia badawczego (np. kwestionariusza), a nie informację gromadzoną od respondenta, np.: założmy, że kwestionariusz składa się z następujących bloków pytań: M – pytania metryczkowe, A – Pytania o społeczność lokalną, B – Pytania o kapitał społeczny – przełożyłoby się to na trzy zmienne typu kategoria:
- zmienna o nazwie M z opisem Metryczka;
 - zmienna o nazwie A z opisem Pytania o społeczność lokalną;
 - zmienna o nazwie B z opisem Pytania o kapitał społeczny.
- 3.3.5.2. typu waga należy użyć, gdy dana zmienna przechowuje wagę obserwacji;
- 3.3.5.3. typ TERC odpowiada kodowi GUS jednostki terytorialnej, w której leżą poszczególne obserwacje (jeśli dane takie są gromadzone w badaniu);
- 3.3.6. skala zmiennej: dychotomiczna/nominalna/porządkowa/interwałowa/ilorazowa;
- 3.3.7. rozmiar zmiennej:
- 3.3.7.1. dla liczb i wag – maksymalna liczba cyfr, z jakich składa się liczba (w wypadku liczb rzeczywistych łącznie cyfr dziesiętnych i ułamkowych);
- 3.3.7.2. dla tekstów – maksymalna liczba znaków;
- 3.3.7.3. dla zmiennej typu TERC – 2, 4 lub 6 w zależności czy przechowywane będą identyfikatory, odpowiednio, województw, powiatów czy gmin;
- 3.3.7.4. dla dat – nie dotyczy (data ma stały rozmiar 19 znaków)
- 3.3.8. dokładność zmiennej
- 3.3.8.1. dla zmiennych rzeczywistych – liczba cyfr ułamkowych w reprezentacji liczby (pusta, jeśli nie dotyczy);
- 3.3.8.2. dla zmiennych typu TERC – rok, z którego pochodzą kody województw/powiatów/gmin;
- 3.3.9. wartość minimalna – tylko dla liczb i dat – minimalna poprawna wartość zmiennej (pusta, jeśli nie dotyczy);
- 3.3.10. wartość maksymalna – tylko dla liczb i dat – maksymalna poprawna wartość zmiennej (pusta, jeśli nie dotyczy);

3.3.11. etykiety wartości – tylko dla zmiennych tekstowych zakodowanych liczbowo – tekst, który w kolejnych liniach zawiera pary kod:etykieta, gdzie kod to kod liczbowy przyporządkowany danej etykiecie (pusta, jeśli nie dotyczy).

3.3.11.1. dla danych pochodzących z ankiet kwestionariuszowych:

3.3.11.1.1. etykiety wartości powinny być takie same, jak teksty odpowiadających im punktów kafeterii danego pytania kwestionariusza;

3.3.11.1.2. opis etykiet wartości musi pokrywać całą kafeterię danego pytania kwestionariusza;

3.3.12. warunek posiadania przez zmienną wartości – tylko dla zmiennych podlegających filtrom – zdanie logiczne opisujące to, kiedy zmienna powinna mieć wartość, o strukturze:

3.3.12.1. w każdym wierszu po jednym warunku logicznym;

3.3.12.2. każdy warunek logiczny postaci: [i/lub] nazwaZmiennej operator wartość, gdzie:

3.3.12.2.1. i/lub na początku warunku oznacza sposób łączenia się danego warunku z warunkami w poprzednich wierszach, przy czym:

3.3.12.2.1.1. obowiązuje matematyczny priorytet warunków i/lub;

3.3.12.2.1.2. człon ten nie jest obecny w pierwszym warunku logicznym;

3.3.12.2.2. dostępne operatory to: =, !=, <, >, <=, >= (równe, różne od, mniejsze, większe, mniejsze bądź równe, większe bądź równe); np.

A1 = 3

i B2 > 5

lub B3 = 9

oznacza warunek: (jeśli na pytanie A1 udzielono odpowiedzi o kodzie 3 oraz na pytanie B2 odpowiedzi o kodzie większym od 5) lub na pytanie B3 udzielono odpowiedzi o kodzie 9

4. Opis struktury zanonimizowanego zbioru danych ilościowych (codebook) powinien opisywać wszystkie zmienne znajdujące się w zanonimizowanym zbiorze danych ilościowych.

5. Zanonimizowany zbiór danych ilościowych powinien zawierać:

5.1. W pierwszym wierszu nazwy zmiennych przechowywanych w poszczególnych kolumnach zbioru (patrz opis struktury zanonimizowanego zbioru danych ilościowych).

- 5.2. W kolejnych wierszach wartości poszczególnych zmiennych dla kolejnych jednostek obserwacji.
- 5.2.1. Jeśli dana zmienna jest typu kategoria, to powinna ona mieć pustą wartość dla wszystkich jednostek obserwacji (patrz 3.3.5.1).
- 5.2.2. Jeśli dana zmienna jest typu data, powinna mieć format: RRRR-MM-DD GG:MM:SS, gdzie RRRR to czterocyfrowy zapis lat, MM to dwucyfrowy zapis numeru miesiąca, DD to dwucyfrowy zapis numeru dnia w miesiącu, GG to dwucyfrowy zapis godzin w formacie 24-godzinnym, MM to dwucyfrowy zapis minut, a SS to dwucyfrowy zapis sekund, np. 2010-02-04 17:06:48, co oznacza drugi lutego 2010 r., sześć minut i 48 sekund po godzinie siedemnastej. Jeśli data pozyskiwana była z mniejszą dokładnością, zapis należy uzupełnić zerami, np. jeśli data pozyskiwana była z dokładnością do dni, zapis powinien mieć postać RRRR-MM-DD 00:00:00.
- 5.2.3. Jeśli dana zmienna nie ma wartości dla danej jednostki obserwacji z uwagi na to, że nie była ona w ogóle gromadzona dla danej jednostki obserwacji (np. wskutek filtru na pytaniu kwestionariusza lub tego, że wykonywany przez danego ucznia zeszyt ćwiczeń nie zawierał danego pytania), powinna zostać jej przypisana wartość:
- 5.2.3.1. dla zmiennych liczbowych, wag, zmiennych typu teryt: maksymalna wartość zmiennej o danym rozmiarze - 2, np. dla zmiennej o rozmiarze 3 cyfr 997;
- 5.2.3.2. dla zmiennych będących datami: 9997-00-00 00:00:00;
- 5.2.3.3. dla zmiennych tekstowych: ciąg znaków NIE DOTYCZY;
- 5.2.4. Jeśli dana zmienna nie ma wartości dla danej jednostki obserwacji z uwagi na odmowę lub nieudzielenie odpowiedzi przez respondenta / ucznia (także gdy uczeń w teście zamkniętym nie zaznaczył żadnej odpowiedzi), a także omyłkowe niezadanie pytania wywiadu kwestionariuszowego przez ankietera:
- 5.2.4.1. dla zmiennych liczbowych, wag, zmiennych typu teryt: maksymalna wartość zmiennej o danym rozmiarze, np. dla zmiennej o rozmiarze 3 cyfr 999;
- 5.2.4.2. dla zmiennych będących datami: 9999-00-00 00:00:00;
- 5.2.4.3. dla zmiennych tekstowych: ciąg znaków ODMOWA ODPOWIEDZI;
- 5.2.5. Jeśli dana zmienna nie ma wartości dla danej jednostki obserwacji z uwagi na to, że respondent / uczeń nie potrafił udzielić odpowiedzi lub w teście zamkniętym zaznaczył więcej niż jedną odpowiedź:
- 5.2.5.1. dla zmiennych liczbowych, wag, zmiennych typu teryt: maksymalna wartość zmiennej o danym rozmiarze - 1, np. dla zmiennej o rozmiarze 3 cyfr 998;
- 5.2.5.2. dla zmiennych będących datami: 9998-00-00 00:00:00;

5.2.5.3. dla zmiennych tekstowych: ciąg znaków TRUDNO POWIEDZIEĆ;

5.2.6. Długości zmiennych liczbowych muszą być dobrane w ten sposób, by opisane powyżej wartości specjalne znajdowały się poza dopuszczalnym zakresem „zwyczajnych” wartości zmiennych.

6. Zanonimizowany zbiór danych ilościowych:

6.1. musi zawierać zmienną stanowiącą unikalny identyfikator obserwacji w zbiorze;

6.2. nie może zawierać zmiennych stanowiących dane osobowe (np. imiona, nazwiska, adres zamieszkania, telefon, itp.);

6.3. musi być w pełni zgodny z opisem swojej struktury.

7. Wszystkie zanonimizowane bazy danych muszą zawierać zmienną będącą identyfikatorem szkoły, której dotyczy dana obserwacja w zbiorze (tzn: w przypadku bazy z ankiet uczniowskich - szkoły, do której uczęszcza badany uczeń; w przypadku bazy z ankiet nauczycielskich – szkoły, w której naucza badany nauczyciel, a w przypadku bazy kart szkół – szkoły, której dotyczy dana karta). Wartość tej zmiennej dla konkretnej szkoły musi być taka sama we wszystkich trzech zbiorach danych).

8. Baza danych z ankiet uczniowskich oraz baza danych z ankiet nauczycielskich muszą też zawierać zmienną będącą identyfikatorem oddziału, którego dotyczy dana obserwacja w zbiorze (tzn: w przypadku bazy z ankiet uczniowskich - oddziału, do którego uczęszcza badany uczeń; w przypadku bazy nauczycielskiej zmienna ta będzie miała wartości tylko dla ankiet od wychowawców badanych oddziałów i będzie dotyczyła oddziału, którego badany jest wychowawcą).

9. Standardy kodowania wybranych zmiennych:

9.1. Jeśli zanonimizowany zbiór danych ilościowych zawiera zmienną określającą gminę, powinna ona być kodowana z użyciem sześciocyfrowego kodu TERC. W opisie struktury zbioru danych taka zmienna powinna mieć typ TERC.

9.2. Jeśli zanonimizowany zbiór danych ilościowych zawiera zmienną identyfikującą podmioty posiadające numer REGON, to powinna ona zawierać numery REGON tych podmiotów.

Zbiory zostały sprawdzone przez zespół badawczy IBE pod kątem zgodności z powyższymi zasadami. W przypadku uwag beneficjent projektu konkursowego poprawił zbiory. Anonimizacji poddano wszelkie informacje, przez które w zbiorach danych można by było zidentyfikować konkretną szkołę lub konkretnego ucznia w szkole. Poniżej zamieszczone zostały informacje charakteryzujące każdy z wytworzonych podstawowych zbiorów. Każdy ze zbiorów został poddany przekształceniu z formatu csv na formaty odczytywalne za pomocą różnych pakietów statystycznych (SPSS, Stata, R).

6.1.1. Zbiór danych z wywiadów z dyrektorami szkół

Nazwa pliku: *MLEZAMiD_I_runda_dyrektor*.

Liczebność: N= 1374.

Identyfikator do łączenia z innymi zbiorami: *KOD_szkoly*.

Nazwy zmiennych: zaczynają się od prefixu „*DYR_*” (z wyjątkiem zmiennej *KOD_szkoly*).

Dodatkowe zmienne dołączone do zbioru: *POZYCJA_szkoly_w_operacie*, *TERYT_szkoly*, *REGON_szkoly*, *ID_SIO_szkoly*.

6.1.2. Zbiór danych z ankiet audytoryjnych z uczniami w szkole

Nazwa pliku: *MLEZAMiD_I_runda_AUDYT*.

Liczebność: N=12248.

Identyfikator do łączenia z innymi zbiorami: *KOD_szkoly* (zbiór z wywiadów z dyrektorami).

Nazwy zmiennych: zaczynają się od prefixu „*AUD_*” (z wyjątkiem zmiennych *KOD_szkoly* oraz *ID_RESP*).

Dodatkowe zmienne dołączone do zbioru: *POZYCJA_szkoly_w_operacie*, *TERYT_szkoly*, *REGON_szkoly*, *ID_SIO_szkoly*.

6.1.3. Zbiór danych z wywiadów z uczniami w szkole (pierwszy kontakt)

Nazwa pliku: *MLEZAiMD_I_runda_CAPI*.

Liczebność: N=10102. Do 7713 jednostek z tego zbioru można dołączyć jednostki ze zbioru absolwentów.

Identyfikator do łączenia z innymi zbiorami: *KOD_szkoly* (zbiór z wywiadów z dyrektorami), *ID_RESP* (zbiór wywiadów absolwentami – drugi kontakt).

Nazwy zmiennych: zaczynają się od prefixu „*UCZ_*” (z wyjątkiem zmiennych *KOD_szkoly* oraz *ID_RESP*).

Dodatkowe zmienne dołączone do zbioru: *POZYCJA_szkoly_w_operacie*, *TERYT_szkoly*, *REGON_szkoly*, *ID_SIO_szkoly*.

6.1.4. Zbiór danych z wywiadów z absolwentami (drugi kontakt)

Nazwa pliku: *MLEZAiMD_I_runda_CAPI_absolwent*.

Liczebność: N=7713 wszystkie jednostki łączą się ze zbiorem wywiadów CAPI z uczniami w szkole (pod identyfikatorze *ID_RESP*).

Identyfikator do łączenia z innymi zbiorami: *KOD_szkoly* (zbiór z wywiadów z dyrektorami), *ID_RESP* (zbiór z wywiadami CAPI z uczniami – pierwszy kontakt).

Nazwy zmiennych: zaczynają się od prefiksu „ABS_” (z wyjątkiem zmiennych *KOD_szkoly* oraz *ID_RESP*).

Dodatkowe zmienne dołączone do zbioru: *POZYCJA_szkoly_w_operacie*, *TERYT_szkoly*, *REGON_szkoly*, *ID_SIO_szkoly*.

6.2. Ważenie poststratyfikacyjne

6.2.1. Przygotowane wagi poststratyfikacyjne

Do zbiorów danych z wynikami I rundy MLEZAIMD przygotowane zostały dwa zestawy wag poststratyfikacyjnych, korygujących różnice w prawdopodobieństwie doboru szkół do próby wynikające ze schematu doboru próby oraz z nierównomierność realizacji terenowej badania pomiędzy województwami i obszarami kształcenia.

Wagi do analiz prowadzonych w podziale ze względu na typ szkoły:

- *waga_audyt* - do analiz zbioru z badania audytoryjnego;
- *waga_capi1f* - do analiz zbioru z 1. fali badania CAPI (tj. badania uczniów);
- *waga_capi2f* - do analiz zbioru z 2. fali badania CAPI (tj. badania absolwentów), lub połączonych zbiorów z 1. i 2. fali badania CAPI;

waga_pop_capi2f - waga do analiz zbioru z 2. fali badania CAPI (tj. badania absolwentów), lub połączonych zbiorów z 1. i 2. fali badania CAPI, wszystkich typów szkół łącznie;

Wagi do analiz uwzględniających wszystkie trzy typy szkół zawodowych łącznie:

- *waga_pop_audyt* – do analiz zbioru z badania audytoryjnego;
- *waga_pop_capi1f* - do analiz zbioru z 1. fali badania CAPI (tj. badania uczniów);
- *waga_pop_capi2f* - do analiz zbioru z 2. fali badania CAPI (tj. badania absolwentów), lub połączonych zbiorów z 1. i 2. fali badania CAPI.

6.2.2. Procedura ważenia

Ze względu na złożoność przeprowadzonego badania procedura obliczania wag poststratyfikacyjnych składała się z kilku etapów. Najpierw badanym przypisywane były

wagi związane z ich prawdopodobieństwem trafienia do próby, wynikającym ze schematu doboru prób do badania. Następnie wagi te były przeliczane, aby uwzględnić modyfikacje schematu doboru próby (w aspekcie wydawania szkół rezerwowych), które wprowadzono w trakcie realizacji terenowej badania. Kolejnym etapem była korekta wartości wag ze względu na poziomi realizacji badania w ramach poszczególnych szkół. Wreszcie, na zakończenie całej procedury, wagi przeliczano z zastosowaniem poststratyfikacji metodą wieńcową (tzw. *raking*), w celu zniwelowania wpływu nierównomierności realizacji terenowej badania pomiędzy województwami i obszarami kształcenia. W kolejnych sekcjach szczegółowo opisane zostały poszczególne etapy zastosowanej procedury.

6.2.3. Konsekwencje schematu doboru próby dla ważenia

Przy obliczaniu wag muszą zostać uwzględnione dwie konsekwencje przyjętego schematu doboru próby i wydawania szkół rezerwowych. Po pierwsze, zawarta w schematach doboru *awersja* do dobierania szkół (szczególnie) dużych musi zostać skompensowana przez przypisanie takim szkołom większych wag, odwrotnie proporcjonalnych do prawdopodobieństwa ich doboru do próby. Po drugie, w sytuacji, gdy wydawane były szkoły rezerwowe, waga jest związana z liczbą uczniów szkoły, która została dobrana do próby podstawowej (na miejsce której wydana została dana szkoła rezerwowa), a nie z liczbą uczniów w szkole, w której zrealizowano badanie. Ogólnie rzecz biorąc, w związku ze wspomnianą wcześniej *awersją* do dobierania szkół (szczególnie) dużych do próby podstawowej, w wyniku wydawania szkół rezerwowych zwiększa się oczekiwana liczba uczniów w próbie zrealizowanej. Skutkuje to z kolei zmniejszeniem wartości wag przy przeliczaniu ich w ten sposób, aby ich suma był równa liczbie obserwacji w zbiorze.

6.2.4. Odstępstwa od pierwotnie założonego schematu

W toku realizacji badania nastąpił szereg odstępstw od pierwotnie założonego schematu doboru próby, wymagających uwzględnienia przy obliczaniu wag. Poniżej zostały one wymienione oraz omówiono konsekwencje, jakie poszczególne odstępstwa miały dla sposobu obliczania wag.

1. Zamiast w ZSZ wylosowanej do próby audytoryjnej, opisanej kodem *AZ1557-3410-0*, badanie audytoryjne zostało zrealizowane w technikum.
 - W zbiorze opisującym zrealizowaną próbę utworzono kopię szkoły o kodzie szkoły *AZ1557-3410-0*, i zmieniono jej typ na technikum.
 - Ponieważ w szkole (ZSZ), która powinna zostać zbadana przeprowadzono wywiady CAPI (por. p. 5. poniżej), pozostawiono w zbiorze ZSZ o tym samym kodzie. Każdej z tych kopii szkoły kodowanej jako *AZ1557-3410-0* wyzerowano odpowiednio liczbę uczniów zbadanych ankietą audytoryjną lub liczbę uczniów zbadanych ankietą CAPI.
2. Kilka zrealizowanych szkół rezerwowych pobrano zamiast tej samej szkoły z próby podstawowej.

- Nominalnie, jeśli zamiast jakiejś szkoły pobrano szkołę rezerwową (lub kolejną szkołę rezerwową), nie powinno w niej zostać zrealizowane badanie (powodem wydania szkoły rezerwowej powinna być niemożliwość realizacji badania w placówce, za którą wydawana była szkoła rezerwowa). W praktyce wykonawca badania niejednokrotnie pobierał szkołę rezerwową *na zapas*, a następnie realizował badanie w placówce, na miejsce której dokonał pobrania z rezerwy.
- Przyjęto, że nie wpływa to na konieczność przeliczania wag (w ramach ogólnej zasady, że waga szkoły rezerwowej jest równa wadze szkół z próby podstawowej, za którą została wydana).

3. W toku badania wyczerpano operat szkół policealnych.

- W związku z wyczerpaniem operatu szkół policealnych dla wszystkich szkół tego typu, w których udało się zrealizować badanie, przyjęto prawdopodobieństwo doboru do próby równe 1.

4. W związku z wyczerpaniem operatu szkół policealnych, w celu osiągnięcia założonej łącznej liczby zbadanych szkół dołosowano do próby dodatkowych 100 zasadniczych szkół zawodowych (ZSZ).

- 100 szkół zostało dobranych w sposób prosty spośród tych zasadniczych szkół zawodowych, które pozostawały w rezerwie na koniec dnia pracy 6.06.2017.
- Prawdopodobieństwo trafienie do próby dla tych szkół wynosiło:

$$100 (1 - p_p) / (n_{ZSZo} - n_{ZSZpp} - n_{ZSZz})$$

gdzie:

p_p prawdopodobieństw wyboru szkoły do próby podstawowej (zdarzenie, które nie zaszło - dlatego we wzorze używa się jego dopełnienia do 1);

n_{ZSZo} = 1145 liczba ZSZ w operacie,

n_{ZSZpp} = 450 liczba ZSZ w próbie podstawowej,

n_{ZSZz} = 171 liczba ZSZ wydanych jako rezerwowe do dn. 6.06.2017 (włącznie).

- Prawdopodobieństwo to było wyraźnie niższe, niż średnie prawdopodobieństwo doboru do próby podstawowej, co wymagało korekty poprzez przypisanie tym szkołom proporcjonalnie wyższych wag.

5. W niektórych szkołach badanie zrealizowano inną techniką, niż tą, która była zakładana na etapie doboru szkoły do próby badawczej.

- Dotyczy to przede wszystkim zrealizowania wywiadów CAPI w szkołach wylosowanych do próby audytoryjnej (lub dobranych na ich miejsce szkół rezerwowych), ale zdarzyło się również, że w szkołach dobranych do próby CAPI

(lub dobranych na ich miejsce szkół rezerwowych) zrealizowano badanie audytoryjne

- Przyjęto, że w ramach szkół dobór uczniów do badania był zgodny z założonym dla danej techniki bez względu na to, czy była ona pierwotnie przewidziana do realizacji w danej szkole (tj. wszyscy uczniowie przy ankiecie audytoryjnej i prosta próba losowa 30 uczniów przy ankiecie CAPI).

6.2.5. Korekta wag ze względu na poziom realizacji badania w szkole

W ramach założonego schematu doboru próby przewidziano stosowanie wag korygujących poziom realizacji badania w szkole (w formie $1/rr$, gdzie rr to odsetek wywiadów zrealizowanych w szkole). Ze względu na bardzo niski poziom realizacji w niektórych placówkach zdecydowano, że **w zbiorze CAPI szkoły, w których zrealizowano tylko 1 lub 2 wywiady zostały wyłączone z badania poprzez przypisanie im wartości wag równych 0**. Uczniów tych szkół wyłączono również z badania absolwentów, przy czym nie stosowano już dodatkowych wyłączeń związanych z poziomem realizacji badania absolwentów. **W zbiorze z badania audytoryjnego nie stosowano wyłączenia szkół z badania, ale maksymalną wartość wagi korygującej poziom realizacji ograniczono do 10**.

6.2.6. Ważenie poststratyfikacyjne

Zgodnie z założeniami określonymi w ekspertyzie dot. doboru próby do badania MLEZAiMD, ważenie poststratyfikacyjne przeprowadzono **oddzielnie dla każdego typu szkoły** metodą ważenia wieńcowego (*raking*) ze względu na:

- obszar kształcenia,
- województwo (w którym znajduje się szkoła).

Ponieważ dla niektórych obserwacji w zbiorach danych obszar kształcenia jest nieznany (uczeń nie podał zawodu), podjęto następujące działania mające na celu uzupełnienie tej informacji:

- Jeśli w tej samej szkole wśród uczniów tej samej płci są zawody należące do tylko jednego obszaru kształcenia, to ten obszar kształcenia został przypisany obserwacji z brakiem danych.
- Jeśli w efekcie powyższej procedury nie udało się uzupełnić obszaru kształcenia, obserwacje z brakami danych były wyłączone z procedury ważenia poststratyfikacyjnego i dołączane do zbioru ponownie, po jej przeprowadzeniu.

Poniżej przedstawione zostały przyjęte populacyjne rozkłady brzegowe ww. zmiennych dla poszczególnych typów szkół (obliczone na podstawie SIO wg stanu na 30.09.2017).

Tabela 6.1. Populacyjne rozkłady brzegowe liczby uczniów w województwach

województwo	Szkoła policealna	Technikum	Zasadnicza szkoła zawodowa
-------------	-------------------	-----------	----------------------------

dolnośląskie	7,3 %	6,6 %	6,6 %
kujawsko-pomorskie	5,8 %	5,8 %	6,7 %
Lubelskie	7,4 %	6,1 %	4,6 %
Lubuskie	1,8 %	2,7 %	3,1 %
Łódzkie	8,0 %	6,0 %	4,8 %
małopolskie	7,8 %	10,1 %	11,2 %
mazowieckie	12,5 %	11,9 %	8,9 %
opolskie	2,0 %	2,7 %	3,0 %
podkarpackie	5,2 %	6,9 %	5,9 %
podlaskie	3,9 %	3,2 %	2,2 %
pomorskie	5,8 %	5,9 %	7,6 %
śląskie	12,1 %	11,6 %	10,5 %
świętokrzyskie	4,2 %	3,7 %	2,9 %
warmińsko-mazurskie	3,5 %	3,7 %	4,6 %
wielkopolskie	8,7 %	9,2 %	13,4 %
zachodniopomorskie	3,9 %	3,9 %	4,2 %
łącznie	100,0 %	100,0 %	100,0 %

W przypadku obszarów kształcenia dla poszczególnych typów szkół dokonano połączenia niektórych z nich na potrzeby przeprowadzenia ważenia poststratyfikacyjnego, zgodnie z założeniami określonymi w ekspertyzie dot. doboru próby do badania. Było to konieczne, aby poradzić sobie z problemem niewielkich liczebności niektórych obszarów kształcenia w ramach próby.

Tabela 6.2. Populacyjne rozkłady brzegowe liczby uczniów w obszarach kształcenia

grupa obszarów kształcenia	Szkoła policealna	Technikum	Zasadnicza szkoła zawodowa
administracyjno-usługowy (A), artystyczny (S)	39,4 %	ndt.	26,8 %
administracyjno-usługowy (A), artystyczny (S), medyczo-społeczny (Z)	ndt.	29,1 %	ndt.
budowlany (B)	ndt.	7,7 %	11,5 %
budowlany (B), elektryczno-elektroniczny (E), mechaniczny i górniczo-hutniczy (M)	7,6 %	ndt.	ndt.
elektryczno-elektroniczny (E)	ndt.	25,6 %	5,8 %
mechaniczny i górniczo-hutniczy (M)	ndt.	10,9 %	34,7 %
rolniczo-leśny z ochroną środowiska (R)	5,5 %	6,3 %	0,6 %
turystyczno-gastronomiczny (T)	1,8 %	20,4 %	20,6 %
medyczo-społeczny (Z)	42,0 %	ndt.	0,0 %
łącznie	100,0 %	100,0 %	100,0 %

Chociaż wyniki badania MLEZAiMD zwykle będą analizowane oddzielnie dla każdego z trzech typów szkół, przygotowane zostały również wagi pozwalające na prowadzenie analiz mających na celu wnioskowanie o populacji wszystkich uczniów ostatnich klas szkół zawodowych z roku szkolnego 2017/2018, bez podziału na typ szkoły. W tym celu przeprowadzono w dodatkowym kroku prostą poststratyfikację ze względu na typ szkoły, przyjmując (na podstawie SIO wg stanu na 30.09.2017) następujące częstości populacyjne: technika: 49,3%, ZSZ: 20,8%, SPolic.: 29,9%.

6.2.7. Unormowanie wag

Wagi wynikające ze schematu doboru próby i wagi poststratyfikacyjne przeznaczone do prowadzenia analiz w podziale na szkoły poszczególnych typów zostały unormowane oddzielnie dla każdego typu szkoły w każdym ze zbiorów danych przez podzielenie przez

średnią wartość tych wag wśród jednostek obserwacji, dla których przyjmowały one wartości większe od 0. Wagi przeznaczone do prowadzenia analiz bez podziału na szkoły poszczególnych typów unormowano oddzielnie w każdym zbiorze danych przez podzielenie przez średnią wartość tych wag wśród jednostek obserwacji, dla których przyjmowały one wartości większe od 0. W tabelach poniżej przedstawione są parametry rozkładów unormowanych wag w poszczególnych zbiorach danych.

Tabela 6.3. Parametry rozkładów wag w zbiorze wyników badania audytoryjnego

typ szkoły	wagi	min.	średnia	maks.
Technikum	przed poststrat.	0,16	1,00	6,48
	po poststrat.	0,19	1,00	11,45
	poststrat. z typem szk.	0,32	1,66	19,03
Zasadnicza szkoła zawodowa	przed poststrat.	0,41	1,00	23,97
	po poststrat.	0,18	1,00	23,59
	poststrat. z typem szk.	0,08	0,43	10,05
Szkoła policealna	przed poststrat.	0,34	1,00	3,43
	po poststrat.	0,13	1,00	6,01
	poststrat. z typem szk.	0,19	1,39	8,36

Tabela 6.4. Parametry rozkładów wag w zbiorze wyników badania CAPI w szkołach

typ szkoły	wagi	min.	średnia	maks.
Technikum	przed poststrat.	0,16	1,00	21,27
	po poststrat.	0,23	1,00	20,00
	poststrat. z typem szk.	0,23	1,00	19,93
Zasadnicza szkoła zawodowa	przed poststrat.	0,15	1,00	22,49
	po poststrat.	0,09	1,00	25,25
	poststrat. z typem szk.	0,05	0,54	13,75
Szkoła policealna	przed poststrat.	0,22	1,00	11,21
	po poststrat.	0,03	1,00	32,78
	poststrat. z typem szk.	0,07	2,43	79,75

Tabela 6.5. Parametry rozkładów wag w zbiorze wyników badania CAPI absolwentów

typ szkoły	wagi	min.	średnia	maks.
Technikum	przed poststrat.	0,16	1,00	21,12
	po poststrat.	0,20	1,00	20,36
	poststrat. z typem szk.	0,19	0,97	19,68
Zasadnicza szkoła zawodowa	przed poststrat.	0,16	1,00	23,99
	po poststrat.	0,09	1,00	20,63
	poststrat. z typem szk.	0,05	0,56	11,65
Szkoła policealna	przed poststrat.	0,23	1,00	11,49
	po poststrat.	0,02	1,00	35,11
	poststrat. z typem szk.	0,05	2,47	86,75

6.3. Zbiory epizodów

6.3.1. Wprowadzenie

Informacje o historii edukacyjnej i zawodowej absolwentów po ukończeniu przez nich szkoły zbierane były w badaniu w specyficzny sposób: respondentów proszono o wymienianie kolejnych szkół (uczelnie), do których starali się dostać i w których

kontynuowali naukę, kolejnych miejsc (i stanowisk) pracy itp. Odnosnie każdego z takich wymienionych *epizodów* historii edukacyjno-zawodowej zadawane były respondentowi pytania mające na celu zebranie bardziej szczegółowych informacji nt. tego, gdzie, jak długo, w jakiej formie i dlaczego się on uczył, pracował lub poszukiwał pracy. W zbiorze danych z wynikami badania (w którym jednego respondenta opisuje jeden wiersz – jest to tzw. *postać szeroka*) zmienne, w których zakodowano odpowiedzi na te pytania, występują w formie, która utrudnia ich analizę: każdy kolejny *epizod* danego typu opisują inne zmienne, mimo że zawierają one analogiczne informacje. Nielatwe jest też śledzenie następstwa czasowego pomiędzy *epizodami* różnych typów.

Rozwiązanie problemu analizy danych odnoszących się do *epizodów* stanowi przekształcenie ich w ten sposób, aby pojedynczy wiersz opisującego je zbioru danych odnosił się do konkretnego *epizodu*. Co za tym idzie, historia edukacyjno-zawodowa jednego respondenta może być w takim zbiorze opisywana przez wiele wierszy (jest to tzw. *postać długa* zbioru danych). Analizowanie danych przechowywanych w takiej formie ma kilka istotnych zalet:

- Pozwala łatwo porównywać ze sobą te same cechy kolejnych *epizodów* tego samego typu, które następują w historii edukacyjno-zawodowej tego samego respondenta.
- Pozwala łatwo zestawiać ze sobą podstawowy zestaw cech odnoszących się do każdego typu *epizodów* (lub przynajmniej kilku różnych typów): typ, czas rozpoczęcia i zakończenia, czy zakończył się on *sukcesem* (w przypadku edukacji).
- Pozwala łatwo śledzić następstwo zdarzeń w historii edukacyjno-zawodowej respondenta.
- Pozwala w łatwy sposób, przy pomocy metod agregacji danych, tworzyć nowe zmienne charakteryzujące w syntetyczny sposób interesujące analitycznie cechy historii edukacyjno-zawodowej poszczególnych respondentów.

Aby umożliwić łatwe prowadzenie analiz wykorzystujących tego typu porównania (przekształcenia), utworzone zostały oddzielne zbiory danych wyników badania opisujące *epizody*.

Na podstawie zbioru z badania CAPI w szkołach (1. fala) przygotowane zostały następujące zbiory *epizodów* (w postaci *długiej*):

- *uczniowie_CAPI-prace* zawierający informacje o epizodach pracy;
- *uczniowie_CAPI-szkolenia* zawierający informacje o różnych odbytych przez respondentów dodatkowych szkoleniach i uzyskanych w ich efekcie uprawnieniach zawodowych.

Na podstawie zbioru z badania absolwentów (2. fala) przygotowane zostały następujące zbiory *epizodów* (w postaci *długiej*):

- *absolwenci-bezrobocie* zawierający informacje o *epizodach* bezrobocia;
- *absolwenci-LOdD* zawierający informacje o *epizodach* nauki w LO dla dorosłych;
- *absolwenci-prace* zawierający informacje o *epizodach* pracy;
- *absolwenci-spolic* zawierający informacje o *epizodach* zdawania do szkół policealnych i nauki w szkołach policealnych;
- *absolwenci-studia* zawierający informacje o *epizodach* zdawania na studia i nauki na studiach;
- *absolwenci-szkolenia* zawierający informacje o różnych uzyskanych przez respondentów dodatkowych szkoleniach i uzyskanych w ich efekcie uprawnieniach zawodowych;
- *absolwenci-epizody_wszystkie*, w którym zebrane zostały informacje ze wszystkich ww. plików.

W analogicznej formie utworzono też zbiory danych opisujący skład gospodarstw domowych respondentów, odpowiednio: *uczniowie_CAPI-gosp_dom.sav/dta* i *absolwenci-gosp_dom.sav/dta*. Choć nie odnoszą się one do historii edukacyjno-zawodowej, to jednak problemy związane z analizą informacji o składzie gospodarstw domowych są takie same, jak w przypadku *epizodów*.

Jako specyficzny rodzaj *epizody* potraktowane zostało też uczęszczanie do szkoły, jako uczeń której respondent trafił do badania. Przy analizach po czasie może być bowiem istotne kiedy (i czy z sukcesem) absolwent ją ukończył, a istnieje w tym zakresie pewne (choć niewielkie) zróżnicowanie.

6.3.2. Różnice w nazwach i etykietach względem oryginalnego zbioru SPSS

Nazwy wszystkich zmiennych w zbiorach *epizodów*, z wyjątkiem zmiennej ID_RESP, przekształcone zostały na pisane małymi literami. Poza tym aspektem nazwy zmiennych w zbiorach *epizodów* zasadniczo odpowiadają nazwom zmiennych z oryginalnego pliku w formacie SPSS, z tym że pozbawione zostały elementów wskazujących na *nr pętli* (w zależności od zmiennej przyrostek tego rodzaju mógł występować albo na końcu nazwy zmiennej, albo nieco wcześniej). Niewielkie modyfikacje nazw, mające na celu umieszczenie analogicznych informacji w tej samej zmiennej, zostały dokonane w odniesieniu do niektórych pytań o pierwszą i ostatnią pracę oraz pytań z bloku U4. Zmiany te zostały opisane szczegółowo w części poświęconej specyfice różnych typów *epizodów*.

Pewne przekształcenia wartości zmiennych wymuszały też ograniczenia techniczne, związane z tym, że przekształcanie danych prowadzone było w środowisku R, podczas gdy wynikowe pliki zapisywane były jako pliki programów SPSS i Stata (tzn. każdy plik zarówno w jednym, jak i w drugim formacie). W związku z tym **wartości powiązane z niektórymi etykietami w plikach w postaci długiej mogą różnić się od tych, jakie są im przypisane w oryginalnym pliku w formacie SPSS**. Typowo dotyczy to kodów

specjalnych opisujących tzw. *braki danych użytkownika* – w praktyce ma to jednak niewielkie znaczenie, gdyż wartości *braków danych użytkownika* przy tworzeniu zbiorów *epizodów* były co do zasady przekodowywane na *systemowe* braki danych. Braki danych w zmiennych tekstowych oznaczone są ciągiem znaków składającym się z pojedynczej kropki: '.' (Stata nie akceptuje pustych ciągów znaków).

Wszystkie zmienne, które mają etykiety wartości "tak"- "nie" zakodowane są w zbiorach *epizodów* wg schematu: 1-tak, 2-nie (w zmiennych sp6h_1 - sp6h_3 i pp6i_1 - pp6i_3 w oryginalnym zbiorze SPSS wartości kodowane były wg schematu 0-nie, 1-tak).

6.3.3. Schemat analizy z wykorzystaniem zbiorów *epizodów*

Typowo zbiory *epizodów* wykorzystywane są w analizie w ten sposób, że:

1. najpierw dokonane zostaną na nich pewne procedury agregacji w ramach respondentów, aby uzyskać wskaźniki opisujące charakterystykę historii edukacyjno-zawodowej respondenta będącej przedmiotem zainteresowania;
2. następnie tak zagregowane wyniki będą dołączane do oryginalnego zbioru danych (w którym jednemu respondentowi odpowiada jeden wiersz) w celu przeprowadzenia dalszych analiz z wykorzystaniem utworzonego wskaźnika.

Należy mieć przy tym na uwadze, że znaczna część respondentów mogła nie mieć żadnego *epizodu* właśnie analizowanego typu i w związku z tym nie występują oni (żaden wiersz, który by ich dotyczył) w odpowiednim zbiorze *epizodów*. W związku z tym, przy dołączaniu wyników zagregowanych do poziomu respondentów na podstawie pliku *epizodów* do oryginalnego zbioru (w którym jednostką obserwacji jest respondent) wielu osobom w oryginalnym zbiorze, obejmującym wszystkich respondentów, mogą zostać przypisane zostaną braki danych (jako wartości wskaźnika obliczonego na zbiorze *epizodów*). Istotne może być więc zakodowanie tych braków danych zaraz po dokonaniu łączenia w sposób, który pozwoli potem na wygodne (i przede wszystkim nieprowadzące do pomyłek, w szczególności związanych z przyjęciem złej podstawy procentowania) obsłużenie ich w prowadzonych dalej analizach.

- Należy przy tym mieć na uwadze, że w plikach w postaci *długiej* wartości *braków danych użytkownika* zostały już typowo zamienione na *systemowe* braki danych (a więc *systemowy* brak danych w połączonym pliku może dla wielu zmiennych równie dobrze wskazywać na to, że respondent nie miał żadnego *epizodu* danego typu, co na to, że taki *epizod* miał, ale nie był w stanie podać o nim interesującej nas informacji).
- W związku z tym dobrą praktyką przy agregacji zbioru danych w postaci *długiej* do poziomu respondentów jest utworzenie w nim takiej zmiennej, która dla wszystkich w zagregowanym zbiorze przyjmie wartość niebędącą brakiem danych i będzie ją można potem wygodnie wykorzystać do stwierdzenia, którzy respondenci w oryginalnym zbiorze danych mieli choć jeden *epizod* danego typu.

Jeśli zagregowany do poziomu respondentów zbiór *epizodów* będzie analizowany bez przyłączania do oryginalnego zbioru, należy mieć na uwadze, że zawiera on typowo tylko

podzbiór spośród wszystkich badanych, co nie zawsze jest pożądaną podstawą procentowania.

6.3.4. Zmienne specjalne w zbiorach *epizodów*

Aby możliwe było łatwiejsze ich analizowanie, we wszystkich zbiorach *epizodów* zostały utworzone specjalne zmienne:

- ID_RESP - identyfikator respondenta, umożliwia łączenie z innymi zbiorami danych z badania;
- typ_epizodu – opisuje typ *epizodu*, przyjmuje jedną z dziesięciu wartości:
 1. 'szkoła objęta badaniem';
 2. 'LO dla dorosłych';
 3. 'zdawanie do SPolic.';
 4. 'SPolic.';
 5. 'zdawanie na studia';
 6. 'studia';
 7. 'szkolenia u2';
 8. 'szkolenia u4';
 9. 'praca';
 10. 'bezrobocie'.
- Epizody 'studia' i 'SPolic.' są specyficzne, gdyż zawierają również informacje dotyczące zdawania na te kierunki studiów (do tych szkół policealnych), na (w) których respondent podjął naukę. W związku z tym, jeśli przedmiotem analizy mają być kierunki (szkoły) na (do) których respondent zdawał, należy w niej uwzględnić odpowiednio zarówno *epizody* typu 'zdawanie na studia', jak i *epizody* typu 'studia' albo zarówno *epizody* typu 'zdawanie do SPolic.', jak i *epizody* typu 'SPolic.'.
- nr - wartość zmiennej wskazuje, jako który z kolei (inaczej mówiąc, w którym obiegu pętli skryptu, czy też w którym wierszu tabeli reprezentującej odpowiedzi na blok pytań o „epizody” danego typu w „papierowej” wersji kwestionariusza) epizod określonego typu respondent wymienił dany epizod;
 - Respondenci niekoniecznie wymieniali *epizody* w zgodzie z ich chronologiczną kolejnością. Jeśli chce się mieć pewność, że *epizody* uszeregowane są w kolejności chronologicznej, należy posłużyć się w pierwszej kolejności zmiennymi czas_rozp lub czas_zakon, a dopiero

w dalszej zmiennej nr (przy czym należy mieć na uwadze, że w zmiennych czas_rozp i czas_zakon występują braki danych, a w przypadku zmiennej czas_zakon należy też obsłużyć sytuacje, gdy *epizod* się nie zakończył).

- Zmienna nr odnosi się łącznie do *epizodów* typu 'zdawanie do SPolic.' i 'SPolic.' (tj. epizody obu tych typów w ramach tej samej osoby posiadają jedną, ciągłą *numerację*) oraz łącznie do *epizodów* typu 'zdawanie na studia' i 'studia', co wiąże się z opisaną wyżej specyfiką *epizodów* typów 'studia' i 'SPolic.'.
 - Dla *epizodów* typu 'LO dla dorosłych' jako wartość zmiennej nr występuje tylko 1, gdyż kwestionariusz nie dopuszczał wymienienia kilku *epizodów* tego typu.
 - W przypadku *epizodów* typu 'praca' **wartość zmiennej nr równa 99** oznacza, że dany *epizod* pracy został *dopowiedziany* w wyniku udzielenia przez respondenta odpowiedzi na pytania PG5.
- czas_rozp - czas rozpoczęcia trwania danego *epizodu*, wyrażony jako **liczba miesięcy od czerwca 2017 r.**;
- zmienna może przyjmować wartości ujemne (typowo zdarza się to w odniesieniu do *epizodów* pracy);
 - nie ma zastosowania dla *epizodów* typu 'szkolenia u2' ani 'szkolenia u4';
 - **jeśli respondent nie podał miesiąca**, w jakim rozpoczął się dany *epizod*, dla *epizodów* typu 'praca' i 'bezrobocie' dokonano imputacji wartości zmiennej czas_rozp (procedura została opisana w dalszej części raportu); dla *epizodów* pozostałych typów zawsze arbitralnie przyjmowano, że był to czerwiec;
 - należy mieć na uwadze, że mimo to zdarzają się *epizody* (do których ma zastosowanie data rozpoczęcia) z brakami danych w tej zmiennej (respondenci czasem nie byli w stanie podać nawet roku);
- czas_zakon - czas zakończenia trwania danego *epizodu*, wyrażony jako **liczba miesięcy od czerwca 2017 r.**;
- zmienna może przyjmować wartości ujemne (typowo zdarza się to w odniesieniu do *epizodów* pracy);
 - nie ma zastosowania dla *epizodów* typu 'zdawanie na studia' i 'zdawanie do SPolic.';
 - **jeśli respondent nie podał miesiąca**, w jakim zakończył się dany *epizod*, dla *epizodów* typu 'praca' i 'bezrobocie' dokonano imputacji wartości zmiennej czas_zakon (procedura została opisana w dalszej części raportu); dla *epizodów* pozostałych typów zawsze arbitralnie przyjmowano, że był to lipiec;

- należy mieć na uwadze, że mimo to zdarzało się, że respondenci nie byli w stanie podać nawet roku i jako wartość zmiennej występuje brak danych;
 - **brak danych w zmiennej czas_zakon może też wskazywać na to, że dany epizod wciąż trwa** - aby to sprawdzić, należy posłużyć się zmienną czy_zakonczony.
- czy_zakonczony - dychotomiczna zmienna wskazująca, czy dany epizod się zakończył, czy w momencie realizacji wywiadu z respondentem wciąż trwał:
 - dla epizodów typu 'LO dla dorosłych', 'studia', 'SPolic.', 'szkolenia u2' i 'szkolenia u4' wartość zmiennej została przekodowana ze zmiennych odpowiednio: zp2c, sp6d, pp6d (jako wskazujące na zakończenie epizodu traktowano odpowiedzi mówiące o uzyskaniu dyplomu, jak i o przerwaniu nauki) oraz u2d;
 - dla epizodów typu 'praca' i 'bezrobocie' wartość zmiennej została zakodowana na podstawie zmiennych odpowiednio: pg2f, i pb1e;
 - przy kodowaniu wartości dla epizodów wszystkich ww. typów, jeśli respondent nie udzielił odpowiedzi na odpowiednie pytanie, przyjmowano, że epizod się zakończył;
 - dla epizodów typu 'zdawanie na studia' i 'zdawanie do SPolic.' wartością zmiennej jest brak danych.
- swiadectwo - dychotomiczna zmienna, w przypadku epizodów typu 'LO dla dorosłych', 'studia', 'SPolic.' 'szkolenia u2' i 'szkolenia u4' wskazująca, czy dany epizod zakończył się uzyskaniem świadectwa (w odniesieniu do trzech pierwszych typów można to utożsamiać z *zakończeniem się sukcesem*, w przypadku szkoleń respondent mógł wymienić kursy, które nie dawały możliwości uzyskania formalnego potwierdzenia ich ukończenia); dla epizodów typu zdawanie na studia i zdawanie do SPolic. zmienna zawsze przyjmuje wartość 2 ('nie'), dla epizodów 'pracy' i 'bezrobocia' wartością jest brak danych;
- czas_rozp_imput - dychotomiczna zmienna opisująca, czy wartość zmiennej czas_rozp została zaimputowana (procedura została opisana w dalszej części raportu); jeśli czas_rozp jest brakiem danych, również czas_rozp_imput jest brakiem danych;
- czas_zakon_imput - dychotomiczna zmienna opisująca, czy wartość zmiennej czas_zakon została zaimputowana (procedura została opisana w dalszej części raportu); jeśli czas_zakon jest brakiem danych, również czas_zakon_imput jest brakiem danych;
- czas_zakon_sz - zmienna opisująca czas zakończenia przez respondenta nauki w szkole, jako uczeń której trafił on do badania - wyrażona na tej samej skali co zmienne czas_rozp i czas_zakon;

- zmienna ta pozwala w łatwy sposób przejść od skali czasu zaczepionej w miesiącu kalendarzowym, w jakiej wyrażone są domyślnie zmienne czas_rozp i czas_zakon do skali zaczepionej w momencie zakończenia nauki w szkole zawodowej (jako absolwent której respondent ma być badany).

6.3.5. Specyfika poszczególnych typów epizodów

6.3.5.1. Epizody dotyczące zdawania do szkół policealnych lub na studia

Na podstawie zmiennych odpowiednio pp3a i sp3a utworzona została zmienna czy_preferowany opisująca to, czy szkoła lub kierunek był preferowany przez respondenta (spośród potencjalnie kilku, do których/na które zdawał) przy pomocy wartości: 1-tak, 2-nie. Jeśli respondent zdawał tylko do jednej szkoły, czy na jedno studia, w zmiennej czy_preferowany przypisana została do danej szkoły/kierunku wartość 1 (tak). Jeśli respondent nie wskazał preferowanego kierunku/szkoły, wartością zmiennej jest brak danych.

6.3.5.2. Informacje o odbytych szkoleniach i zdobytych uprawnieniach

Pytania o zdobyte uprawnienia i odbyte szkolenia zadawane były w kwestionariuszu w dwóch blokach: U2 i U4, z których pierwszy w założeniach odnosił się do szkoleń kończących się uzyskaniem oficjalnych certyfikatów (względnie uzyskania takich certyfikatów bez dedykowanego szkolenia), a drugi do pozostałych szkoleń. Trudno jednak powiedzieć, w jakim stopniu respondenci odpowiadając trzymali się tego podziału.

Pytania w obu blokach były analogiczne, poza tym, że w bloku U2 wystąpiły dwa pytania dotyczące samego egzaminu, których nie było w bloku U4. W związku z tym nazwy zmiennych opisujących odpowiedzi na pytania z bloku U4 zostały zmienione na zaczynające się od *u2* (zamiast *u4*), tak aby epizody obu tych typów (tj. *szkolenia u2* i *szkolenia u4*) można było analizować łącznie. Dodatkowo nazwę zmiennej *u4g* zmieniono na *u2h*, gdyż w przypadku tych dwóch zmiennych w kwestionariuszu przypisano inne sufiksy literowe do analogicznych pytań.

Blok pytań, w którym opisany został dany epizod można zidentyfikować na podstawie zmiennej typ_epizodu.

W zbiorze z badania uczniów zadawane były tylko pytania z bloku U2. Również w przypadku tego zbioru dokonano zmiany nazwy jednej zmiennej: *u2g* zmieniono na *u2h*, aby uspołnić schemat nazewniczy ze zbiorem z badania absolwentów (pytanie oznaczone jako U2g w kwestionariuszu absolwenta nie występowało w kwestionariuszu ucznia - jego dodanie do ankiety doprowadziło do przesunięcia w oznaczeniu pytań).

6.3.5.3. Epizody pracy

Zmienna *pg1b* podaje kod pracodawcy, który w sytuacji, gdy respondent miał wiele epizodów pracy, pozwala stwierdzić, czy była to praca u różnych, czy u tego samego pracodawcy. Kody (numery) pracodawców pozwalają ich identyfikować tylko w ramach rekordów opisujących tego samego respondenta.

W zbiorze utworzonym na podstawie wyników 1. fali badania (*uczniowie_CAPI-prace*) występuje w odniesieniu do tej informacji pewne ograniczenie: dla *epizodów* pracy, które zostały *dopowiedziane* (w odpowiedzi na pytanie, czy respondent nie przypomniał sobie żadnych dodatkowych *epizodów* pracy), nazwa pracodawcy nie została przez wykonawcę badania zakodowana w sposób analogiczny do tego, jak postąpiono w zasadniczej pętli pytań o pracę. W związku z tym konieczne było przyjęcie arbitralnego założenia, że *epizod* ten dotyczy innego pracodawcy, niż wcześniej wymienieni. W takim przypadku jako kod pracodawcy wpisana została wartość 99.

Zakres informacji o *epizodach* pracy różni się w zależności od tego, czy był to pierwszy, ostatni, czy któryś inny *epizod* pracy. W przypadku pierwszych i ostatnich *epizodów* dołączono bowiem do zbioru informacje o odpowiedziach na pytania odpowiednio PI0-PI5 oraz PO1-PO11. W przypadku pytań P1-P5 i PO1-PO5, które zadawane były w tej samej formie, utworzone zostały w zbiorze wspólne zmienne opisujące udzielone na nie odpowiedzi, zastępując w nazwach oryginalnych zmiennych przedrostki "pi" i "po" wspólnym przedrostkiem "pio".

Chociaż pytania PO7-PO11 były zadawane wszystkim respondentom, którzy wymienili choć jeden *epizod* pracy, to jednak nie miały one sensu dla tych, którzy w momencie realizacji wywiadu nie pracowali (przez pomyłkę nie dodano filtru w ankiecie). Dla tych osób odpowiedzi udzielone na te pytania zostały zastąpione brakami danych.

Dokonane zostało *domknięcie* odpowiedzi na pytanie o formę zatrudnienia, podanych przez respondentów w formie opisowej (oryginalne odpowiedzi znajdują się zmiennej ABS_pg2h_in):

- odpowiedzi dające się zidentyfikować jako pasujące wprost do podanej w pytaniu kafeferii przekodowano odpowiednio;
- "kontrakt" został potraktowany jako samozatrudnienie (zmiennej ABS_pg2h przypisano wartość 5);
- różne typy staży zostały dołączone do stażu z UP (zmiennej ABS_pg2h przypisano wartość 7);
- pracę u rodziców na polu potraktowano jako pracę *u siebie w gospodarstwie* (zmiennej ABS_pg2g przypisano wartość 7, zmiennej ABS_pg2h brak danych).

Kwestionariusz w badaniu uczniów (1. fala badania) zawierał zawężony zestaw pytań dot. pracy w stosunku do kwestionariusza w badaniu absolwentów (2. fala). W związku z tym zestaw zmiennych dostępnych w zbiorze *uczniowie_CAPI-prace* jest odpowiednio mniejszy.

6.3.5.4. Epizody bezrobocia

W odpowiedziach na pytanie o przyczynę niepodejmowania pracy ciężę wymienianą w części "inne" przekodowana na "opiekę nad dzieckiem" (zmiennej pb1g przypisano wartość 2).

6.3.6. Imputacja wartości zmiennych czas_rozp i czas_zakon

W analizie historii edukacyjno-zawodowych absolwentów zasadnicze znaczenie ma możliwość uporządkowania w czasie kolejnych *epizodów*, jak też określenie długości ich trwania. Niestety niekiedy respondenci mieli trudności z podaniem, kiedy opisywany przez nich *epizod* się zaczął lub skończył. Ze względu na istotność tych informacji zastosowane zostały procedury mające na celu zaimputowanie powstałych w konsekwencji braków danych w zmiennych *czas_rozp* i *czas_zakon*. Zakres wykorzystanych przy tym technik rozciąga się od arbitralnego przypisania najbardziej prawdopodobnych wartości do imputacji z wykorzystaniem przewidywania na podstawie modeli regresji. W zestawieniach poniżej przedstawione zostały liczby braków danych przed i po przeprowadzeniu imputacji.

Tabela 6.6. Liczba i częstość występowania braków danych w zmiennych *czas_rozp* i *czas_zakon* dla *epizodów* różnych typów przed przeprowadzeniem imputacji

<i>typ epizodu</i>	liczba <i>epizodów</i>	b.d. <i>czas_rozp</i> liczba <i>epizodów</i>	%	b.d. <i>czas_zakon</i> liczba <i>epizodów</i>	%
Bezrobocie	2490	105	4,2	48	1,9
LO dla dorosłych	978	15	1,5	8	0,8
Praca	6987	328	4,7	125	1,8
SPolic.	386	9	2,3	1	0,3
Studia	1564	15	1,0	13	0,8
szkolenia u2	3222	ndt.	ndt.	956	29,7
szkolenia u4	267	ndt.	ndt.	67	25,1

Tabela 6.7. Liczba i częstość występowania braków danych w zmiennych *czas_rozp* i *czas_zakon* dla *epizodów* różnych typów po przeprowadzeniu imputacji

<i>typ epizodu</i>	liczba <i>epizodów</i>	b.d. <i>czas_rozp</i> liczba <i>epizodów</i>	%	b.d. <i>czas_zakon</i> liczba <i>epizodów</i>	%
bezrobocie	2490	83	3,3	39	1,6
LO dla dorosłych	978	6	0,6	5	0,5
praca	6987	147	2,1	78	1,1
SPolic.	386	2	0,5	1	0,3
studia	1564	12	0,8	11	0,7
szkolenia u2	3222	ndt.	ndt.	192	6,0
szkolenia u4	267	ndt.	ndt.	18	6,7

6.3.6.1. Arbitralne przypisanie wartości

W przypadku wszystkich *epizodów* innych typów, niż 'praca' i 'bezrobocie', jeśli respondent podał rok, ale nie podał miesiąca rozpoczęcia/zakończenia *epizodu*, arbitralnie przyjmowano, że rozpoczął się on w lipcu lub skończył w czerwcu.

W przypadku *epizodów* nauki sprawdzano, czy tak przypisana wartość zmiennej *czas_zakon* nie okazuje się mniejsza, niż wartość zmiennej *czas_rozp* (może do tego dojść, jeśli respondent *wypadł* ze szkoły jeszcze w tym samym roku, w którym rozpoczął w niej naukę). Jeśli tak było, jako wartość *czas_zakon* arbitralnie przypisywano liczbę o 1 większą od wartości *czas_rozp*.

Tak samo postąpiono z datami rozpoczęcia/zakończenia epizodów pracy i bezrobocia, co do których respondenci zadeklarowali, że zaczęły/skończyły się one w latach innych niż 2017 lub 2018 (ze względu na niewielką ogólną liczbę *epizodów* obejmujących taki wcześniejszy okres nie ma możliwości sensownie zaimputować tych czasów na podstawie modelu regresji). Nie zdarzyło się, aby epizod o przypisanej w ten sposób wartości zmiennej *czas_rozp* miał niepełną informację dotyczącą czasu zakończenia epizodu (tzn. tak przypisane wartości zmiennej *czas_rozp* nie zostały ani razu wykorzystane w procedurze opisanej w następnej sekcji).

Przypisane w omówiony wyżej sposób wartości oznaczone są w zmiennych *czas_rozp_imput* i *czas_zakon_imput* jako zaimputowane.

6.3.6.2. Uzupełnianie wartości zmiennych *czas_rozp* i *czas_zakon* w epizodach pracy i bezrobocia nieoznaczone jako imputacja (tylko dane z 2. fali)

Opisane poniżej działania przeprowadzono tylko w odniesieniu do danych z badania absolwentów (2. fali), gdyż w 1 fali badania (badanie uczniów) nie zbierano danych o długości trwania epizodu.

Jeśli respondent dla danego *epizodu*:

- nie podał informacji koniecznych do obliczenia wartości zmiennej *czas_rozp*, ale dało się ustalić wartość zmiennej *czas_zakon* lub
- nie podał informacji koniecznych do obliczenia wartości zmiennej *czas_zakon*, ale dało się ustalić wartość zmiennej *czas_rozp*,

a także udzielił odpowiedzi na pytanie o szacunkową długość trwania epizodu (zmiennie *pg2x*, *pb1x*), to wartość *brakującej* zmiennej obliczano odpowiednio dodając lub odejmując długość trwania epizodu do/od znanego momentu rozpoczęcia lub zakończenia. Obliczone w ten sposób wartości **nie są** w zmiennych *czas_rozp_imput* i *czas_zakon_imput* oznaczane jako zaimputowane.

6.3.6.3. Imputacja na podstawie modeli regresji (tylko dane z 2. fali)

Imputacji brakujących wartości zmiennych *czas_rozp* i *czas_zakon* dla *epizodów* typu 'praca' i 'bezrobocie' dokonano przy użyciu modeli regresji liniowej. Imputację tą metodą przeprowadzono tylko w sytuacji, gdy brakowało informacji o miesiącu, ale była podana informacja o roku rozpoczęcia/zakończenia *epizodu*. Modele estymowane były oddzielnie dla *epizodów* pracy i bezrobocia. W procedurze imputacji (zarówno przy estymacji modeli, jak i przy samym imputowaniu braków danych) brano pod uwagę wyłącznie epizody, które rozpoczęły się w 2017 r. lub później.

Zmienna *czas_rozp*

- Zmienną zależną w modelu była *czas_rozp*.
- Jako predyktory wykorzystano:

- 1) typ szkoły (dwie zmienne *dummy* zakodowane na podstawie zmiennej typ_szkoły);
- 2) płeć respondenta (zmienna m1);
- 3) formę pracy lub rozróżnienie na bezrobotnych (poszukujących pracy) i biernych zawodowo (przekodowane na zestaw zmiennych *dummy*, przy czym brak danych uwzględniono przy przekodowaniu jako oddzielną wartość);
- 4) rok, w którym rozpoczął się *epizod* (traktowany jako zmienna kategorialna - zakodowany przy pomocy zmiennej *dummy*);
- 5) liczbę wszystkich *epizodów* pracy danego respondenta;
- 6) liczbę wszystkich *epizodów* bezrobocia (lub nieaktywności zawodowej) danego respondenta;
- 7) informację o tym, czy w (dowolnym momencie) 2017 roku respondent brał udział w edukacji formalnej, a jeśli tak, to w jakiej formie (4 zmienne *dummy*);
- 8) informację o tym, czy w (dowolnym momencie) 2018 roku respondent brał udział w edukacji formalnej, a jeśli tak, to w jakiej formie (4 zmienne *dummy*);
- 9) informację o tym, w którym miesiącu został przeprowadzony wywiad z respondentem (zmienna r5s2 traktowana jako zmienna ciągła);
- 10) interakcje pomiędzy 1) a 2), 4), 5), 6) i 9);
- 11) interakcje pomiędzy 2) a 3), 4), 5), i 6).

Model osiągał następujące statystyki dopasowania (na zbiorze epizodów, dla których znane były wartości zmiennej czas_rozp):

- Dla *epizodów* typu 'praca': $R^2 = 0,68$; odsetek poprawnych klasyfikacji = 16,6%.
- Dla *epizodów* typu 'bezrobocie': $R^2 = 0,74$; odsetek poprawnych klasyfikacji = 38,7%.

Przy dokonywaniu imputacji przewidywania wynikające z modelu zaokrąglano do najbliższych liczb całkowitych (tak aby zmienna czas_rozp również po imputacji przyjmowała tylko wartości całkowite). Ww. odsetki poprawnych klasyfikacji zostały obliczone na podstawie tak zaokrąglonych wartości (ale wartość R^2 podano dla niezaokrąglonych przewidywań).

Po dokonaniu imputacji dokonano sprawdzenia:

- Czy zaimputowana wartość zmiennej zgadza się ze znanym rokiem rozpoczęcia *epizodu*.

- Czy zaimputowana wartość zmiennej nie wskazuje, że *epizod* rozpoczął się później, niż się zakończył (dla *epizodów*, dla których znana była wartość zmiennej *czas_zakon*, ale wartość zmiennej *czas_rozp* była imputowana).

Nie stwierdzono jednak wystąpienia żadnego z ww. problemów.

Skuteczność modelu imputacji, zwłaszcza w odniesieniu do epizodów pracy, pozostawia nieco do życzenia. W praktyce okazał się on mieć tendencje do plasowania większości przewidywanych czasów rozpoczęcia w okolicach wartości 2-3 oraz 8-9 (w zależności od roku, w którym zaczął się epizod). Jednakże biorąc pod uwagę, że zastosowanie do imputacji modelu regresji pozwala włączyć do analizy blisko 100 dodatkowych *epizodów* pracy, niedokładność ta wydaje się możliwa do zaakceptowania.

Zmienna *czas_zakon*

- Zmienną zależną w modelu była długość trwania *epizodu* (tj. różnica pomiędzy *czas_zakon* a *czas_rozp*).
- Przy imputacji wartość zmiennej *czas_zakon* określano dodając do wartości zmiennej *czas_rozp* wartość przewidywania wynikającą z modelu (zaokrągloną do najbliższej liczby całkowitej).
- Jako predyktory wykorzystano:
 - 1) typ szkoły (dwie zmienne *dummy* zakodowane na podstawie zmiennej *typ_szkoły*);
 - 2) płeć respondenta (zmienna *m1*);
 - 3) formę pracy lub rozróżnienie na bezrobotnych (poszukujących pracy) i biernych zawodowo - p. opis zmiennych *praca* i *bezrobocie* w sekcji *Zbiór z danymi w postaci osobo-miesiący* (przekodowane na zestaw zmiennych *dummy*, przy czym brak danych uwzględniono przy przekodowaniu jako *legalną* wartość);
 - 4) rok, w którym rozpoczął się *epizod* (traktowany jako zmienna kategoryalna - zakodowany przy pomocy zmiennej *dummy*);
 - 5) rok, w którym zakończył się *epizod* (traktowany jako zmienna kategoryalna - zakodowany przy pomocy zmiennej *dummy*);
 - 6) interakcję 4) i 5);
 - 7) liczbę wszystkich *epizodów* pracy danego respondenta;
 - 8) liczbę wszystkich *epizodów* bezrobocia (lub nieaktywności zawodowej) danego respondenta;
 - 9) informację o tym, w którym miesiącu został przeprowadzony wywiad z respondentem (zmienna *r5s2* traktowana jako zmienna ciągła);

- 10) interakcje pomiędzy 1) a 2), 4), 7, i 8);
- 11) interakcje pomiędzy 2) a 3)-8);
- 12) interakcje pomiędzy 3) a 5);
- 13) interakcje pomiędzy 4) a 7) i 8);
- 14) interakcje pomiędzy 5) a 7), 8) i 9).

Model osiągał następujące statystyki dopasowania (na zbiorze epizodów, dla których znane były wartości zmiennej *czas_zakon*):

- Dla *epizodów* typu 'praca': $R^2 = 0,66$; odsetek poprawnych klasyfikacji = 27,4%.
- Dla *epizodów* typu 'bezrobocie': $R^2 = 0,71$; odsetek poprawnych klasyfikacji = 27,5%.

Przy dokonywaniu imputacji przewidywania wynikające z modelu zaokrąglano do najbliższych liczb całkowitych (tak aby zmienna *czas_zakon* również po imputacji przyjmowała tylko wartości całkowite). Ww. odsetki poprawnych klasyfikacji zostały obliczone na podstawie tak zaokrąglonych wartości (ale wartość R^2 podano dla niezaokrąglonych przewidywań).

Po dokonaniu imputacji dokonano sprawdzenia:

- Czy zaimputowana wartość zmiennej zgadza się ze znanym rokiem zakończenia *epizodu*.
- Czy zaimputowana wartość zmiennej nie wskazuje, że *epizod* zakończył się wcześniej, niż się rozpoczął.

Nie stwierdzono jednak wystąpienia żadnego z ww. problemów.

Niestety trzeba stwierdzić, że skuteczność modelu imputacji pozostawia nieco do życzenia. W praktyce model nie jest w stanie skutecznie przewidywać bardzo długich (można by powiedzieć *nietypowo długich*) *epizodów*. Jednocześnie jednak w odniesieniu do tych trwających od 0 do 9 miesięcy (a więc typowych, które mogą się pojawić w zbiorze) jest co prawda nie bardzo dokładny, ale dający zadowalające rezultaty.

6.4. Zbiory osobo-miesięcy

6.4.1. Wprowadzenie

Pewne metody analizy historii edukacyjno-zawodowej absolwentów wymagają posłużenia się danymi w jeszcze innej formie: zbioru, w którym poszczególne wiersze (jednostki obserwacji) opisują cechy respondentów w kolejnych momentach czasu, np. w kolejnych miesiącach. Zbiory w tej postaci wykorzystywane są np. w tzw. *analizach przeżycia*, a po pewnych przekształceniach również w tzw. *analizach sekwencji*. Poniżej opisany został sposób przygotowania zbioru danych w tej formie na podstawie zbioru *epizodów*.

6.4.2. Struktura zbioru osobo-miesiący

Dla każdego respondenta w zbiorze zawarte są rekordy opisujące każdy kolejny miesiąc od września 2016 r. (włącznie) do miesiąca poprzedzający ten, w którym nastąpiła realizacja wywiadu z danym absolwentem. Analizowany okres można oczywiście łatwo zawęzić, odfiltrując obserwacje na podstawie zmiennej czas.

W zbiorze zawarte zostały następujące zmienne opisujące cechy respondentów stałe w czasie:

- ID_RESP - identyfikator respondenta, umożliwia łączenie z innymi zbiorami danych z badania;
- typ_szkoły - typ szkoły, jako uczeń której respondent został zakwalifikowany do badania;
- r5s2 - miesiąc, w którym został przeprowadzony wywiad z respondentem;
- f6 - zmienna opisująca powód ew. niezyskania świadectwa szkoły, jako uczeń której respondent został zakwalifikowany do badania (jeśli badany nie uzyskał świadectwa; jeśli respondent ukończył tę szkołę, zmiennej przypisany jest brak danych);
- f7 - rok uzyskania świadectwa szkoły, jako uczeń której respondent został zakwalifikowany do badania (jeśli respondent nie uzyskał świadectwa tej szkoły, zmiennej przypisany jest brak danych);
- m1 - płeć respondenta;
- m2 - rok urodzenia respondenta;
- m3 - klasa wielkości miejscowości zamieszkania respondenta w momencie przeprowadzania wywiadu.

W zbiorze zawarte zostały następujące zmienne opisujące statusy respondentów w poszczególnych miesiącach:

- data - identyfikator miesiąca w formie *czytelnej*;
- czas - identyfikator miesiąca w formie liczby - liczba miesięcy, jaka upłynęła od czerwca 2017 r. (a więc *domyślnego* momentu ukończenia szkoły, jako absolwent której respondent został zakwalifikowany do badania) do miesiąca, który opisuje dany rekord w zbiorze;
- status - zmienna kodująca w syntetycznej formie status respondenta: wartości zmiennych (kolejno) praca, nauka i bezrobocie *sklejone* ze sobą w jeden ciąg znaków (składający się z trzech cyfr: cyfra setek koduje pracę, dziesiątek naukę, a jedności bezrobocie);
- wartość zmiennej status równa '999' ma specjalne znaczenie - p. następna sekcja;

- praca - zmienna kodująca status zatrudnienia;
- nauka - zmienna kodująca status uczestnictwa w edukacji formalnej;
- bezrobocie - zmienna kodująca, czy respondent deklaruje się jako pozostający bez pracy;
- praca_a_bezrobocie - zmienna opisująca ew. występowanie w danym miesiącu (dla danego respondenta) *konfliktów* pomiędzy statusem zatrudnienia a statusem bezrobocia, fakt skorygowania takiego *konfliktu* (p. sekcja *Procedura kodowania statusu edukacyjno-zawodowego*) lub jego przyczynę;
- korekta_ciąglosc_nauki - zmienna pozwala zidentyfikować rekordy, w których status uczestnictwa w nauce formalnej został skorygowany ze względu na stwierdzenie ciągłości nauki (p. sekcja *Procedura kodowania statusu edukacyjno-zawodowego*);
- imput_praca - zmienna wskazująca, że status zatrudnienia w danym miesiącu (dla danego respondenta) został zakodowany w oparciu o zaimputowaną wartość zmiennej *czas_rozp* lub zmiennej *czas_zakon*;
- imput_nauka - zmienna wskazująca, że status uczestnictwa w nauce formalnej w danym miesiącu (dla danego respondenta) został zakodowany w oparciu o zaimputowaną wartość zmiennej *czas_rozp* lub zmiennej *czas_zakon*;
- imput_bezrobocie - zmienna wskazująca, że status bezrobocia w danym miesiącu (dla danego respondenta) został zakodowany w oparciu o zaimputowaną wartość zmiennej *czas_rozp* lub zmiennej *czas_zakon*.

6.4.3. Procedura kodowania statusu edukacyjno-zawodowego

Kodowanie statusu pracy nauki i bezrobocia przebiegało w ten sposób, że dla każdego respondenta, dla każdego miesiąca uwzględnionego w zbiorze wyszukiwane były wszystkie *epizody*, które obejmowały ten miesiąc. Następnie kodowano ww. statusy, biorąc pod uwagę zarówno cechy *epizodów*, które obejmowały ten miesiąc, jak i informacje o tym, że danego miesiąca nie obejmowały żadne *epizody* danego typu (w szczególności wnioskowanie o statusie osoby biernej zawodowo oparte jest na takim pośrednim wnioskowaniu na podstawie informacji o tym, że badany nie deklaruje, aby pracował, uczył się, lub poszukiwał pracy). Poniżej opisane zostały wszystkie stosowane w tym celu reguły oraz przypisywane wartości statusów.

6.4.3.1. Reguły klasyfikacji na podstawie *epizodów* obejmujących poszczególne miesiące

Przy kodowaniu pominięte zostały *epizody*, dla których nieznaną (już po imputacji) był czas rozpoczęcia *epizodu* lub *epizod* był zakończony (*czy_zakonczony* równe 1) i nieznaną (już po imputacji) był czas zakończenia *epizod*. – w związku z tym nie wzięto pod uwagę 278 *epizodów* (1,4% wszystkich odnotowanych *epizodów*).

praca - zmienną zakodowano na podstawie zmiennych *ABS_pg2g* i *ABS_pg2h*:

reguła	zakodowana wartość
ABS_pg2h równe 1 (ABS_pg2g równe 1, 2 lub 3)	1 (zatrudniony na umowę o pracę na czas określony)
ABS_pg2h równe 2 (ABS_pg2g równe 1, 2 lub 3)	2 (zatrudniony na umowę o pracę na czas nieokreślony)
ABS_pg2h równe 3 (ABS_pg2g równe 1, 2 lub 3)	3 (zatrudniony przez agencję pracy tymczasowej)
ABS_pg2h równe 4 (ABS_pg2g równe 1, 2 lub 3)	4 (zatrudniony na umowie cywilnoprawnej)
ABS_pg2h równe 5 (ABS_pg2g równe 1, 2 lub 3)	5 (samozatrudniony – „praca u kogoś”)
ABS_pg2g równe 4	6 (prowadzi własną działalność – „praca u siebie”);
ABS_pg2g równe 5	7 (prowadzi własne gosp. rolne)
ABS_pg2h równe 7 lub 8 (ABS_pg2g równe 1, 2 lub 3)	8 (odbywa staż lub praktykę absolwencką)
ABS_pg2h równe 6 (ABS_pg2g równe 1, 2 lub 3)	9 (zatrudniony bez umowy, „na czarno”)
wszystkie pozostałe przypadki, gdy wystąpił epizod pracy	10

nauka - zmienną zakodowano na podstawie zmiennej typ_epizodu:

reguła	zakodowana wartość
typ_epizodu równe 'studia'	2
typ_epizodu równe 'SPolic.'	3
wszystkie miesiące od wrześnie 2016 r. do czerwca 2017 r. (wartości zmiennej czas od -9 do 0)	1

wartość 1 oznacza, że respondent uczył się w tym czasie w szkole, jako uczeń której został zakwalifikowany do badania; jeśli nie ukończył jej w czerwcu 2017 r., status ten mógł występować również w następnych miesiącach;

bezrobocie - zmienną zakodowano na podstawie zmiennej pb1f:

reguła	zakodowana wartość
pb1f równe 1	1 (bezrobotny, poszukuje pracy)
pb1f równe 2	2 (bierny zawodowo)
wszystkie pozostałe przypadki, gdy wystąpił epizod bezrobocia	3

Uwzględnienie ciągłości nauki (przez okres wakacji)

Jeśli stwierdzono, że respondent:

- studiował w październiku danego roku i jednocześnie uczył się w szkole dowolnego typu w maju tego samego roku,
- lub uczył się w szkole policealnej we wrześniu danego roku i jednocześnie uczył się w szkole dowolnego typu w maju tego samego roku,

wszystkie miesiące (włącznie) od czerwca do odpowiednio września lub sierpnia, w których respondent nie zadeklarował, że się uczył, oznaczane zostały jako okres odpowiednio studiowania (nauka równe 2) lub nauki w szkole policealnej (nauka równe 3).

Kodowanie na podstawie niewystępowania epizodów obejmujących dany miesiąc

praca - jeśli respondent nie wymienił żadnych *epizodów* pracy obejmujących dany miesiąc, zmiennej praca przypisywana była wartość 0;

nauka - jeśli respondent nie wymienił żadnych *epizodów* nauki obejmujących dany miesiąc, zmiennej nauka przypisywana była wartość 0 (z wyłączeniem miesięcy, dla których status ten zmieniono w wyniku opisanej powyżej procedury uwzględniania ciągłości nauki);

bezrobocie:

- jeśli respondent nie wymienił żadnego *epizodu* bezrobocia obejmującego dany miesiąc i jednocześnie choć jedna zmienna spośród praca i nauka przyjmowała w tym miesiącu dla respondenta wartość inną niż 0, zmiennej bezrobocie przypisywano wartość 0 (*nie bezrobotny*);
- jeśli respondent nie wymienił żadnego *epizodu* bezrobocia obejmującego dany miesiąc i jednocześnie zarówno zmienna praca, jak i zmienna nauka przyjmowały wartość 0 (tj. respondent nie pracował, ani się nie uczył), przyjmowano, że respondent był w danym miesiącu bierny zawodowo i zmiennej bezrobocie przypisywano wartość 2;
- wystąpienie takiej sytuacji kodowane jest specjalną wartością zmiennej status: '999'.

6.4.4. Relacje pomiędzy statusem pracy i nauki a statusem bezrobocia

W ramach przyjętego schematu kodowania nie narzucano wzajemnego wykluczania się *niezerowych* (tj. wskazujących na aktywność zawodową lub edukacyjną) wartości zmiennych praca i nauka z *niezerowymi* wartościami zmiennej bezrobocie (wskazującymi na pozostawaniu bez zatrudnienia). W związku z występowaniem w zbiorze rekordów, w których statusy wskazują jednocześnie na bezrobocie (lub nieaktywność zawodową) i aktywność zawodową lub naukę, należy rozważyć, że po dużej części są to sytuacje, które mogły mieć miejsce. W szczególności:

- Jeśli przejście od bezrobocia do pracy lub odwrotnie odbyło się w środku miesiąca, respondent istotnie był w danym miesiącu zarówno pracujący jak i bezrobotny (wiersze w zbiorze, w których najprawdopodobniej mamy do czynienia z taką sytuacją wyróżnione są przez wartość 'miesiąc graniczny' zmiennej praca_a_bezrobocie).
- Jeśli respondent pracował *na czarno*, mógł jednocześnie być formalnie bezrobotny (choć w zbiorze nie odnotowujemy tego typu deklaracji).
- Postrzeganie się w kategoriach osoby bezrobotnej może być dla części respondentów oderwane od faktu kontynuowania lub nie nauki.

W sytuacjach, gdy zidentyfikowano *konflikt* statusu pracy i bezrobocia, niedający się zinterpretować w kategoriach *miesiąca granicznego*, starano się je jeszcze rozstrzygnąć, biorąc pod uwagę, że niektóre statusy zostały przypisane na podstawie imputowanych

wartości zmiennych `czas_rozp` i `czas_zakon` i w związku z tym można je uznać za mniej wiarygodne.

- Jeśli w danym miesiącu (dla danego respondenta) status pracy wskazujący na aktywność zawodową został zakodowany na podstawie imputowanych wartości zmiennych `czas_rozp` i `czas_zakon`, a status bezrobocia wskazujący na brak pracy został zakodowany w oparciu o deklaracje respondenta dot. miesiąca rozpoczęcia lub zakończenia *epizodu* (bezrobocia), status zatrudnienia był zmieniany na brak zatrudnienia. Sytuacje takie oznaczone są w zbiorze wartością 'skorygowano pracę' zmiennej `praca_a_bezrobocie`.
- Jeśli w danym miesiącu (dla danego respondenta) status bezrobocia wskazujący na brak pracy został zakodowany na podstawie imputowanych wartości zmiennych `czas_rozp` i `czas_zakon`, a status zatrudnienia wskazujący aktywność zawodową został zakodowany w oparciu o deklaracje respondenta dot. miesiąca rozpoczęcia lub zakończenia *epizodu* (pracy), status bezrobocia był zmieniany na 'nie bezrobotny'. Sytuacje takie oznaczone są w zbiorze wartością 'skorygowano bezrobocie' zmiennej `praca_a_bezrobocie`.
- W innych przypadkach nie dało się dokonać korekt. Są one oznaczone są w zbiorze wartościami 'sprzeczne deklaracje resp.' lub 'sprzeczne wyniki imputacji' zmiennej `praca_a_bezrobocie`, w zależności od tego, czy oba statusy zostały określone na podstawie informacji o miesiącu rozpoczęcia lub zakończenia *epizodu* podanych przez respondenta, czy też oba zostały przypisane na podstawie imputowanych wartości zmiennych `czas_rozp` i `czas_zakon`. Przypadków takich jest na szczęście w zbiorze zaledwie 36 (na blisko 15 tys.).

W ogólności nie mamy niestety gwarancji, że respondenci odpowiadali w sposób spójny. W ramach wywiadu nie mieliśmy bowiem możliwości wyłapywania ew. konfliktów w deklaracjach na bieżąco i zmuszania respondentów do ich rozstrzygnięcia.

7. Wskaźniki w raportach dla szkół

7.1. Procedura obliczania wskaźników

Przygotowanie wskaźników charakteryzujących losy edukacyjno-zawodowe absolwentów poszczególnych szkół jest złożonym procesem, w którym można wyodrębnić następujące etapy:

1. Obróbka i integracja danych dotyczących indywidualnych absolwentów.
2. Obliczenie wartości wskaźników charakteryzujących sytuację edukacyjno-zawodową indywidualnych absolwentów.
3. Obliczenie wartości wskaźników charakteryzujących sytuację edukacyjno-zawodową grup absolwentów.

Pierwszy z ww. etapów docelowo będzie obejmował eksport danych z rejestrów i baz danych administracyjnych. W efekcie uzyskiwane będą zbiory danych, w których jednego absolwenta może opisywać wiele rekordów (*osobo-epizodów* lub *osobo-miesiący* – por. rozdziały 6.3 i 6.4). W ramach pierwszej rundy monitoringu zrealizowane tu działania polegały na kontroli jakości i obróbce zbiorów z wynikami przeprowadzonych badań sondażowych, w celu nadania im właśnie takiej formy, umożliwiającej łatwe obliczenie wskaźników charakteryzujących poszczególnych absolwentów. Działania te zostały szczegółowo opisane we wcześniejszych częściach tego raportu. Opracowane procedury będą mogły być, po niewielkich adaptacjach, zastosowane również w kolejnych rundach monitoringu, w których informacje o losach absolwentów będą zbierane przy pomocy metod sondażowych. Z drugiej strony, uzyskanie dostępu do danych z rejestrów i baz danych administracyjnych będzie wymagało opracowania nowych rozwiązań zapewniających realizację tego etapu prac.

Drugi etap obliczania wskaźników obejmuje przekształcenia, w wyniku których otrzymywany jest zbiór, w którym każdemu z absolwentów odpowiada tylko jeden rekord, a poszczególne zmienne-wskaźniki opisują istotne aspekty jego sytuacji edukacyjno-zawodowej w konkretnym momencie. Zestaw wskaźników obliczonych na podstawie badań zrealizowanych w ramach pierwszej rundy monitoringu opisany został szczegółowo w kolejnym podrozdziale. Można przyjąć, że w kolejnych rundach monitoringu nie powinien on ulegać znacznym zmianom. Wykorzystanie w przyszłości danych z rejestrów i baz danych administracyjnych będzie wiązało się z jednej strony z koniecznością ograniczenia zestawu używanych wskaźników, jako że część z obecnie użytych odnosi się do opinii badanych wyrażonych w badaniu sondażowym, niedostępnych jednak w danych administracyjnych. Z drugiej strony możliwe powinno być rozszerzenie zestawu wskaźników o charakteryzujące te same, co obecnie opisywane, charakterystyki sytuacji absolwentów, ale w dłuższej perspektywie czasu.

Ostatni etap przygotowywania wskaźników wykorzystywanych w raportach dla szkół to agregacja wskaźników charakteryzujących sytuację poszczególnych absolwentów, tak aby w syntetyczny sposób charakteryzowały one sytuację kończących daną szkołę. Zestaw przygotowywanych wskaźników oraz sposób ich obliczania na podstawie

wskaźników indywidualnych opisany został w dalszej części tego rozdziału. Aby ułatwić interpretację wartości wskaźników prezentowanych w raportach zdecydowano się przy tym obliczać je nie tylko na poziomie szkół, ale również szerszych grup – w pierwszej rundzie monitoringu jest to typ szkoły zawodowej. Mają one za zadanie służyć jako *grupy porównawcze*, pomagając ustalić punkt odniesienia do oceny wartości wskaźnika w poszczególnych szkołach. W tym celu wartość każdego wskaźnika w odpowiedniej *grupie porównawczej* jest prezentowana w raportach szkół obok wartości wskaźnika w danej placówce.

Funkcje pozwalające obliczyć wartości wskaźników na podstawie danych z pierwszej rundy monitoringu zostały zaimplementowane w pakiecie *MLASZdane*, działającym w środowisku R.

7.2. Wskaźniki obliczane na poziomie indywidualnym

Gwiazdka przy nazwie wskaźnika oznacza, że jest on możliwy do obliczenia wyłącznie na podstawie wyników badań sondażowych.

7.2.1. Informacje o szkole i o uczniu

- **SZK_kod** – identyfikator szkoły, jednoznacznie identyfikujący placówkę,
- **SZK_typ** – typ szkoły zawodowej: „Szkoła policealna”, „Technikum” lub „Zasadnicza Szkoła Zawodowa”.
- **SZK_powiat** – kod TERYT powiatu, na terenie którego pracuje szkoła,
- **UCZ_zawod** – kod Klasyfikacji Zawodów Szkolnictwa Zawodowego zawodu, którego absolwent uczył się w badanej szkole,
- **UCZ_plec** – ‘K’ lub ‘M’,

7.2.2. Wskaźniki dotyczące wyników egzaminów

- **matura_zdana** – zmienna binarna tworzona na podstawie odpowiedzi na pytanie F8 (oraz informacji o typie szkoły zawodowej):
 - 1 – absolwent zdał maturę w roku ukończenia technikum,
 - 0 – absolwent nie zdał matury w roku ukończenia technikum (obejmuje również sytuacje, gdy absolwent nie podszedł do matury),
 - brak danych – absolwenci szkół zawodowych innych, niż technikum lub brak informacji o tym, czy absolwent zdał maturę.
- **egz_zaw_zdany** – zmienna binarna tworzona na podstawie odpowiedzi na pytania F9 i F10:
 - 1 – absolwent uzyskał dyplom potwierdzający kwalifikacje zawodowe w zawodzie, którego uczył się w badanej szkole, nie później, niż w roku ukończenia tej szkoły,

- 0 – absolwent nie uzyskał dyplomu potwierdzającego kwalifikacje zawodowe w zawodzie, którego uczył się w badanej szkole, nie później, niż w roku ukończenia tej szkoły,
- brak danych – brak informacji o tym, czy absolwent uzyskał dyplomu potwierdzającego kwalifikacje zawodowe w zawodzie, którego uczył się w badanej szkole.

7.2.3. Wskaźniki opisujące pierwszą pracę podjętą przez absolwenta w ciągu 9 miesięcy od miesiąca ukończenia szkoły

Przy obliczaniu wskaźników nie były brane pod uwagę prace świadczone bez umowy.

- **pio1_pierwsza*** – ocena zgodności pierwszej pracy wykonywanej po ukończeniu szkoły z zawodem, którego absolwent uczył się w badanej szkole; tworzona na podstawie odpowiedzi na pytanie PI1:
 - 1 – „tak”,
 - 2 – „nie, ale wymagała podobnej wiedzy i umiejętności”,
 - 3 – „nie, wymagała zupełnie innej wiedzy i umiejętności”,
 - 4 – „w tej pracy nie było ważne, jaką mam wiedzę i umiejętności”,
 - brak danych – absolwent nie podjął pracy po ukończeniu szkoły lub nie odpowiedział na pytanie PI1.
- **pio4_pierwsza** – miesięczne zarobki netto w pierwszej pracy wykonywanej po ukończeniu szkoły; tworzona na podstawie odpowiedzi na pytanie PI4:
 - liczba [PLN],
 - brak danych - absolwent nie pracował po ukończeniu szkoły lub nie odpowiedział na pytanie PI4.
- **pi5_pierwsza*** – powód zakończenia pracy, która była pierwszą pracą wykonywaną po ukończeniu szkoły; tworzona na podstawie odpowiedzi na pytanie PI5:
 - 1 – „nie przedłużono ze mną umowy o pracę”,
 - 2 – „odszedłem/odeszłam z własnej inicjatywy”,
 - 3 – „stanowisko pracy, firma została zlikwidowana”,
 - brak danych – absolwent nie pracował po ukończeniu szkoły lub pracował po ukończeniu szkoły tylko w jednym miejscu i pracował tam wciąż w momencie badania lub nie odpowiedział na pytanie PI5.

- **praca_czas_rozp_pierwsza** – liczba miesięcy, od czerwca w roku ukończenia szkoły do momentu podjęcia pierwszej pracy wykonywanej po ukończeniu szkoły; tworzona na podstawie odpowiedzi na pytania PG2C i PG2D:
 - liczba [miesiące],
 - brak danych – absolwent nie pracował po ukończeniu szkoły lub pracował, ale kontynuował pracę, którą rozpoczął przed ukończeniem szkoły lub nie odpowiedział na pytanie PG2D.

- **praca_przed_ukonczeniem_szkoly_pierwsza** – zmienna logiczna; tworzona na podstawie odpowiedzi na pytania PG2C i PG2D:
 - PRAWDA – pierwsza praca, w jakiej pracował absolwent po ukończeniu szkoły została przez niego podjęta przed ukończeniem tej szkoły,
 - FAŁSZ – pierwsza praca, w jakiej pracował absolwent po ukończeniu szkoły została przez niego podjęta po ukończeniu tej szkoły,
 - brak danych – absolwent nie pracował po ukończeniu szkoły lub nie odpowiedział na pytanie PG2D.

- **pg2gh.1_pierwsza** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2H:
 - PRAWDA – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w formie „umowa o pracę na czas określony (w tym umowa na okres próbny, na zastępstwo)”,
 - FAŁSZ – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w innej formie,
 - brak danych – absolwent nie pracował po ukończeniu szkoły.

- **pg2gh.2_pierwsza** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2H:
 - PRAWDA – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w formie „umowa o pracę na czas nieokreślony”,
 - FAŁSZ – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w innej formie,
 - brak danych – absolwent nie pracował po ukończeniu szkoły.

- **pg2gh.3_pierwsza** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2H:
 - PRAWDA – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w formie „praca przez Agencję Pracy Tymczasowej”,

- FAŁSZ – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w innej formie,
 - brak danych – absolwent nie pracował po ukończeniu szkoły.
- **pg2gh.4_pierwsza** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2H:
 - PRAWDA – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w formie „umowa zlecenia/umowa o dzieło”,
 - FAŁSZ – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w innej formie,
 - brak danych – absolwent nie pracował po ukończeniu szkoły.
- **pg2gh.5_pierwsza*** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2H:
 - PRAWDA – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w formie „samozatrudnienie” (i w pytaniu PG2G absolwent zadeklarował, że jest to „praca u kogoś”),
 - FAŁSZ – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w innej formie,
 - brak danych – absolwent nie pracował po ukończeniu szkoły.
- **pg2gh.6_pierwsza** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2G:
 - PRAWDA – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w formie „Praca u siebie: w firmie”.
 - FAŁSZ – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w innej formie,
 - brak danych – absolwent nie pracował po ukończeniu szkoły.
- **pg2gh.7_pierwsza** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2G:
 - PRAWDA – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w formie „Praca u siebie: w gospodarstwie rolnym”,
 - FAŁSZ – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w innej formie,
 - brak danych – absolwent nie pracował po ukończeniu szkoły.

- **pg2gh.8_pierwsza** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2H:
 - PRAWDA – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w formie „staż z urzędu pracy” lub „umowa o praktykę absolwencką”,
 - FAŁSZ – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w innej formie,
 - brak danych – absolwent nie pracował po ukończeniu.

- **pg2gh.9_pierwsza** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytania PG2G i PG2H:
 - PRAWDA – w pierwszej pracy wykonywanej po ukończeniu szkoły absolwent był zatrudniony w formie innej, niż wymienione w opisanych wcześniej zmiennych lub odmówił odpowiedzi na pytanie o formę zatrudnienia,
 - FAŁSZ – absolwent pracował po ukończeniu szkoły i udzielił odpowiedzi na pytania PG2G i PG2H (to drugie, jeśli się do niego odnosiło) dotyczące formy zatrudnienia w pierwszej pracy wykonywanej po ukończeniu szkoły,
 - brak danych – absolwent nie pracował po ukończeniu.

- **pg2i.1_pierwsza*** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2I:
 - PRAWDA – pierwszą pracę wykonywaną po ukończeniu szkoły absolwent wykonywał „w Polsce, w tej samej miejscowości, w której mieszkałem/mieszkałam mając 14 lat”,
 - FAŁSZ – pierwszą pracę wykonywaną po ukończeniu szkoły absolwent wykonywał w innym miejscu,
 - brak danych – absolwent nie pracował po ukończeniu szkoły.

- **pg2i.2_pierwsza*** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2I:
 - PRAWDA – pierwszą pracę wykonywaną po ukończeniu szkoły absolwent wykonywał „w Polsce, w innej miejscowości niż ta, w której mieszkałem/mieszkałam mając 14 lat”,
 - FAŁSZ – pierwszą pracę wykonywaną po ukończeniu szkoły absolwent wykonywał w innym miejscu,
 - brak danych – absolwent nie pracował po ukończeniu szkoły.

- **pg2i.3_pierwsza** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2I:

- PRAWDA – pierwszą pracę wykonywaną po ukończeniu szkoły absolwent wykonywał „za granicą”,
 - FAŁSZ – pierwszą pracę wykonywaną po ukończeniu szkoły absolwent wykonywał w innym miejscu,
 - brak danych – absolwent nie pracował po ukończeniu szkoły.
- **pg2i.9_pierwsza** – zmienna logiczna; utworzona na podstawie odpowiedzi na pytanie PG2I:
- PRAWDA – absolwent pracował po ukończeniu szkoły, ale nie udzielił odpowiedzi na pytanie PG2I dotyczące miejsca pracy w pierwszej pracy wykonywanej po ukończeniu szkoły,
 - FAŁSZ – absolwent pracował po ukończeniu szkoły i udzielił odpowiedzi na pytanie PG2I dotyczące miejsca pracy w pierwszej pracy wykonywanej po ukończeniu szkoły,
 - brak danych – absolwent nie pracował po ukończeniu szkoły.

7.2.4. Wskaźniki opisujące ostatnią pracę wykonywaną przez absolwenta w ciągu 9 miesięcy od miesiąca ukończenia szkoły

Przy obliczaniu wskaźników nie były brane pod uwagę prace świadczone bez umowy.

- **pio1_ostatnia***, **pio4_ostatnia** – przygotowywane analogicznie, jak wskaźniki odnoszące się do pierwszej pracy; tworzone na podstawie odpowiedzi na pytanie PO1 i PO4.
- **po5_ostatnia*** – ocena, w jakim stopniu ostatnia praca wykonywana w ciągu 9 miesięcy od miesiąca ukończenia szkoły odpowiada oczekiwaniom absolwenta (ogólnie rzecz biorąc); tworzona na podstawie odpowiedzi na pytanie PO5:
 - liczba całkowita od 1 do 6 (skala ocen szkolnych: wyższa wartość wskazuje na lepszą ocenę);
 - brak danych - absolwent nie pracował po ukończeniu szkoły lub nie odpowiedział na pytanie PO5.
- **po6_1_ostatnia*** – ocena, w jakim stopniu ostatnia praca wykonywana w ciągu 9 miesięcy od miesiąca ukończenia szkoły odpowiada oczekiwaniom absolwenta pod względem poziomu zarobków; tworzona na podstawie odpowiedzi na pytanie PO6.a):
 - liczba całkowita od 1 do 6 (skala ocen szkolnych: wyższa wartość wskazuje na lepszą ocenę);
 - brak danych - absolwent nie pracował po ukończeniu szkoły lub nie odpowiedział na pytanie PO6.a).

- **po6_2_ostatnia*** – ocena, w jakim stopniu ostatnia praca wykonywana w ciągu 9 miesięcy od miesiąca ukończenia szkoły odpowiada oczekiwaniom absolwenta pod względem atmosfery w pracy; tworzona na podstawie odpowiedzi na pytanie PO6.b):
 - liczba całkowita od 1 do 6 (skala ocen szkolnych: wyższa wartość wskazuje na lepszą ocenę);
 - brak danych - absolwent nie pracował po ukończeniu szkoły lub nie odpowiedział na pytanie PO6.b).

- **po6_3_ostatnia*** – ocena, w jakim stopniu ostatnia praca wykonywana w ciągu 9 miesięcy od miesiąca ukończenia szkoły odpowiada oczekiwaniom absolwenta pod względem ilości obowiązków; tworzona na podstawie odpowiedzi na pytanie PO6.c):
 - liczba całkowita od 1 do 6 (skala ocen szkolnych: wyższa wartość wskazuje na lepszą ocenę);
 - brak danych - absolwent nie pracował po ukończeniu szkoły lub nie odpowiedział na pytanie PO6.c).

- **po6_4_ostatnia*** – ocena, w jakim stopniu ostatnia praca wykonywana w ciągu 9 miesięcy od miesiąca ukończenia szkoły odpowiada oczekiwaniom absolwenta pod względem możliwości awansu; tworzona na podstawie odpowiedzi na pytanie PO6.d):
 - liczba całkowita od 1 do 6 (skala ocen szkolnych: wyższa wartość wskazuje na lepszą ocenę);
 - brak danych - absolwent nie pracował po ukończeniu szkoły lub nie odpowiedział na pytanie PO6.d).

- **po6_5_ostatnia*** – ocena, w jakim stopniu ostatnia praca wykonywana w ciągu 9 miesięcy od miesiąca ukończenia szkoły odpowiada oczekiwaniom absolwenta pod względem elastyczności godzin pracy; tworzona na podstawie odpowiedzi na pytanie PO6.e):
 - liczba całkowita od 1 do 6 (skala ocen szkolnych: wyższa wartość wskazuje na lepszą ocenę);
 - brak danych - absolwent nie pracował po ukończeniu szkoły lub nie odpowiedział na pytanie PO6.e).

- **po6_6_ostatnia*** – ocena, w jakim stopniu ostatnia praca wykonywana w ciągu 9 miesięcy od miesiąca ukończenia szkoły odpowiada oczekiwaniom absolwenta pod względem możliwości rozwoju kariery; tworzona na podstawie odpowiedzi na pytanie PO6.f):
 - liczba całkowita od 1 do 6 (skala ocen szkolnych: wyższa wartość wskazuje na lepszą ocenę);
 - brak danych - absolwent nie pracował po ukończeniu szkoły lub nie odpowiedział na pytanie PO6.f).

- **pg2gh.1_ostatnia, pg2gh.2_ostatnia, pg2gh.3_ostatnia, pg2gh.4_ostatnia, pg2gh.5_ostatnia*, pg2gh.6_ostatnia, pg2gh.7_ostatnia, pg2gh.8_ostatnia, pg2gh.9_ostatnia** – przygotowywane analogicznie, jak wskaźniki odnoszące się do pierwszej pracy.
- **pg2i.1_ostatnia*, pg2i.2_ostatnia*, pg2i.3_ostatnia, pg2i.9_ostatnia** – przygotowywane analogicznie, jak wskaźniki odnoszące się do pierwszej pracy.

7.2.5. Wskaźniki opisujące sytuację zawodową absolwenta w szóstym miesiącu od miesiąca ukończenia szkoły

Rozpatrując sytuację absolwenta w konkretnym miesiącu brane były pod uwagę wszystkie jego miejsca pracy, w których był zatrudniony w tym miesiącu. Wskaźniki pokazują, czy choć jedna z tych prac była danego typu lub wykonywana w danym miejscu.

Przy obliczaniu wskaźników nie były brane pod uwagę prace świadczone bez umowy.

- **pg2gh.1_6m, pg2gh.2_6m, pg2gh.3_6m, pg2gh.4_6m, pg2gh.5_6m*, pg2gh.6_6m, pg2gh.7_6m, pg2gh.8_6m, pg2gh.9_6m** – przygotowywane analogicznie, jak wskaźniki odnoszące się do pierwszej pracy, z tym że:
 - PRAWDA – **przynajmniej jedna** z prac wykonywanych w szóstym miesiącu od miesiąca ukończenia szkoły była pracą w danej formie,
 - FAŁSZ – **żadna** z prac wykonywanych w szóstym miesiącu od miesiąca ukończenia szkoły nie była pracą w danej formie,
 - brak danych – absolwent nie pracował w szóstym miesiącu od miesiąca ukończeniu szkoły.
- **pg2i.1_6m*, pg2i.2_6m*, pg2i.3_6m, pg2i.9_6m** – przygotowywane analogicznie, jak wskaźniki odnoszące się do pierwszej pracy, z tym że:
 - PRAWDA – **przynajmniej jedną** z prac wykonywanych w szóstym miesiącu od miesiąca ukończenia szkoły absolwent wykonywaną w danej lokalizacji,
 - FAŁSZ – **żadna** z prac wykonywanych w szóstym miesiącu od miesiąca ukończenia szkoły nie była pracą wykonywaną w danej lokalizacji,
 - brak danych – absolwent nie pracował w szóstym miesiącu od miesiąca ukończeniu szkoły.

7.2.6. Wskaźniki opisujące sytuację zawodową absolwenta w dziewiątym miesiącu od miesiąca ukończenia szkoły

Rozpatrując sytuację absolwenta w konkretnym miesiącu brane były pod uwagę wszystkie jego miejsca pracy, w których był zatrudniony w tym miesiącu. Wskaźniki pokazują, czy choć jedna z tych prac była danego typu lub wykonywana w danym miejscu.

Przy obliczaniu wskaźników nie były brane pod uwagę prace świadczone bez umowy.

- **pg2gh.1_9m, pg2gh.2_9m, pg2gh.3_9m, pg2gh.4_9m, pg2gh.5_9m*, pg2gh.6_9m, pg2gh.7_9m, pg2gh.8_9m, pg2gh.9_9m** – przygotowywane analogicznie, jak wskaźniki opisujące sytuację w szóstym miesiącu od ukończenia szkoły.
- **pg2i.1_9m*, pg2i.2_9m*, pg2i.3_9m, pg2i.9_9m** – przygotowywane analogicznie, jak wskaźniki opisujące sytuację w szóstym miesiącu od ukończenia szkoły.

7.2.7. Wskaźniki opisujące, przez jaką część okresu 9 miesięcy od miesiąca ukończenia szkoły absolwent pracował lub był bezrobotny

Przy obliczaniu wskaźników brano pod uwagę tylko naukę formalną (w dowolnym trybie).

Przy obliczaniu tych wskaźników (odmiennie niż w procedurze przygotowywania zbioru osobo-miesiący, omówionej w rozdziale 6.4) jako czasu nauki **nie** traktowano okresu, jaki upłynął pomiędzy zakończeniem nauki w badanej szkole a ew. rozpoczęciem nauki w kolejnej (okresu *wakacji*).

Przy obliczaniu wskaźników nie były brane pod uwagę prace świadczone bez umowy.

- **praca_czas_gdy_bez_nauki_p9m** – liczba miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent pracował (w dowolnej formie, wykluczając zatrudnienie bez umowy) i jednocześnie się nie uczył podzielona przez liczbę miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent się nie uczył:
 - liczba z przedziału [0; 1];
 - brak danych – absolwent przez cały okres się uczył.
- **praca_czas_gdy_nauka_p9m** – liczba miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent pracował (w dowolnej formie, wykluczając zatrudnienie bez umowy) i jednocześnie się uczył podzielona przez liczbę miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent się uczył:
 - liczba z przedziału [0; 1]
 - brak danych – absolwent przez cały okres się nie uczył.
- **praca_czas_p9m** – liczba miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent pracował (w dowolnej formie, wykluczając zatrudnienie bez umowy) podzielona przez 9 (tj. długość okresu):
 - liczba z przedziału [0; 1].
- **praca_czas_gdy_bez_nauki_uop_p9m** – liczba miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent pracował w formie umowy o pracę (na czas określony, lub na czas nieokreślony) i jednocześnie się nie uczył podzielona przez

liczbę miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent się nie uczył:

- liczba z przedziału [0; 1];
 - brak danych – absolwent przez cały okres się uczył.
- **praca_czas_gdy_nauka_uop_p9m** – liczba miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent pracował w formie umowy o pracę (na czas określony, lub na czas nieokreślony) i jednocześnie się uczył podzielona przez liczbę miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent się uczył:
- liczba z przedziału [0; 1];
 - brak danych – absolwent przez cały okres się nie uczył.
- **praca_czas_uop_p9m** – liczba miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent pracował w formie umowy o pracę (na czas określony, lub na czas nieokreślony) podzielona przez 9 (tj. długość okresu):
- liczba z przedziału [0; 1].
- **bezrobocie_czas_gdy_bez_nauki_p9m** – liczba miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent był zarejestrowany jako bezrobotny w urzędzie pracy i jednocześnie się nie uczył podzielona przez liczbę miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent się nie uczył:
- liczba z przedziału [0; 1];
 - brak danych – absolwent przez cały okres się uczył.
- **bezrobocie_czas_gdy_nauka_p9m** – liczba miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent był zarejestrowany jako bezrobotny w urzędzie pracy i jednocześnie się uczył podzielona przez liczbę miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent się uczył:
- liczba z przedziału [0; 1];
 - brak danych – absolwent przez cały okres się nie uczył.
- **bezrobocie_czas_p9m** – liczba miesięcy, w ciągu okresu 9 miesięcy od miesiąca ukończenia szkoły, w czasie których absolwent był zarejestrowany jako bezrobotny w urzędzie pracy podzielona przez 9 (tj. długość okresu):
- liczba z przedziału [0; 1].

7.2.8. Wskaźniki opisujące, czy w kolejnych miesiącach po ukończeniu szkoły absolwent był bezrobotny

Przy obliczaniu wskaźników brano pod uwagę tylko naukę formalną (w dowolnym trybie).

Przy obliczaniu tych wskaźników (odmiennie niż w procedurze przygotowywania zbioru osobo-miesiący, omówionej w rozdziale 6.4) jako czasu nauki **nie** traktowano okresu, jaki upłynął pomiędzy zakończeniem nauki w badanej szkole a ew. rozpoczęciem nauki w kolejnej (okresu *wakacji*).

Przy obliczaniu wskaźników nie były brane pod uwagę prace świadczone bez umowy.

- **bezrobocie_1m, bezrobocie_2m, bezrobocie_3m, bezrobocie_4m, bezrobocie_5m, bezrobocie_6m, bezrobocie_7m, bezrobocie_8m, bezrobocie_9m** – zmienne binarne opisujące status bezrobocia absolwentów w kolejnych miesiącach licząc od miesiąca ukończenia szkoły; tworzone na podstawie odpowiedzi na pytania PB1H, PG2G i PG2H oraz obliczonych czasów rozpoczęcia i zakończenia *epizodów* nauki;
 - 1 – absolwent w danym miesiącu (licząc od miesiąca ukończenia szkoły) był zarejestrowany jako bezrobotny w urzędzie pracy;
 - 0 – absolwent w danym miesiącu (licząc od miesiąca ukończenia szkoły) pracował lub uczył się (i nie był zarejestrowany jako bezrobotny w urzędzie pracy);
 - brak danych – absolwent w danym miesiącu (licząc od miesiąca ukończenia szkoły) nie był zarejestrowany jako bezrobotny w urzędzie pracy, nie pracował, ani się nie uczył.

7.2.9. Wskaźniki opisujące, czy w danych miesiącach po ukończeniu szkoły absolwent uczył się

Przy obliczaniu wskaźników brano pod uwagę tylko naukę formalną (w dowolnym trybie).

- **nauka_6m, nauka_9m** – zmienne tekstowe opisujące, w jakiej szkole uczył się absolwent w odpowiednio szóstym lub dziewiątym miesiącu po ukończeniu szkoły; tworzone na podstawie odpowiedzi na pytania ZP2A, ZP2B, ZP2C, ZP2F, ZP2G, SP6C1, SP6C2, SP6D, SP6E1, SP6E2, SP6H, PP6C1, PP6C2, PP6D, PP6F i PP6F2:
 - „LO dla dorosłych”;
 - „szkoła policealna”;
 - „studia niestacjonarne”;
 - „studia stacjonarne”;
 - brak danych – absolwent w danym miesiącu (licząc od miesiąca ukończenia szkoły) nie był uczniem żadnej szkoły ww. typów.

Jeśli absolwent w danym miesiącu uczył się w szkołach kilku różnych typów, wskaźnikowi przypisywano wartość odpowiadającą temu z nich, który na powyższej liście wymieniany jest najniżej.

- **nauka_platna_6m, nauka_platna_9m** – zmienne binarne opisujące, czy nauka w szkole opisywanej we wskaźnikach odpowiednio **nauka_6m** i **nauka_9m** była odpłatna; tworzone na podstawie odpowiedzi na pytania ZP2I, SP6F i PP6G:
 - 1 – nauka była odpłatna;
 - 0 – nauka była darmowa;
 - brak danych – absolwent w danym miesiącu (licząc od miesiąca ukończenia szkoły) nie był uczniem żadnej szkoły spośród typów wymienionych w opisie wskaźników **nauka_6m** i **nauka_9m**.

7.2.10. Wskaźniki opisujące studia wybierane przez absolwentów

- **studia_kierunek_pierwsze** – zmienna tekstowa tworzona na podstawie odpowiedzi na pytanie SP6B:
 - ciąg znaków – nazwa pierwszego kierunku studiów, które podjął absolwent po ukończeniu szkoły (w okresie 9 miesięcy od miesiąca ukończenia szkoły);
 - brak danych – absolwent nie podjął studiów w badanym okresie lub nie podał nazwy kierunku studiów.
- **studia_uczelnia_pierwsze** – zmienna tekstowa tworzona na podstawie odpowiedzi na pytanie SP6B:
 - ciąg znaków – nazwa uczelni prowadzącej pierwszy kierunek studiów, które podjął absolwent po ukończeniu szkoły (w okresie 9 miesięcy od miesiąca ukończenia szkoły);
 - brak danych – absolwent nie podjął studiów w badanym okresie lub nie podał nazwy uczelni.
- **studia_bezplatne_pierwsze**, – zmienna binarna opisująca, czy pierwsze studia, które podjął absolwent po ukończeniu szkoły (w okresie 9 miesięcy od miesiąca ukończenia szkoły) były odpłatne; tworzona na podstawie odpowiedzi na pytanie SP6F:
 - 1 – studia były darmowe;
 - 0 – studia były odpłatne;
 - brak danych – absolwent nie podjął studiów w badanym okresie lub nie podał nazwy uczelni, lub nie udzielił odpowiedzi na pytanie SP6F.

- **studia_tryb_pierwsze**, – tryb kształcenia na pierwszych studiach, które podjął absolwent po ukończeniu szkoły (w okresie 9 miesięcy od miesiąca ukończenia szkoły); tworzona na podstawie odpowiedzi na pytanie SP6H:
 - 1 – studia stacjonarne (dzienne);
 - 2 – studia niestacjonarne (wieczorowe);
 - 3 – studia niestacjonarne (weekendowe, zaoczne);
 - brak danych – absolwent nie podjął studiów w badanym okresie lub nie podał nazwy uczelni, lub nie udzielił odpowiedzi na pytanie SP6H.

7.2.11. Wskaźniki opisujące status edukacyjno-zawodowy absolwenta w kolejnych miesiącach od ukończenia szkoły

Przy obliczaniu wskaźników brano pod uwagę tylko naukę formalną (w dowolnym trybie).

Przy obliczaniu tych wskaźników (odmiennie niż w procedurze przygotowywania zbioru osobo-miesiący, omówionej w rozdziale 6.4) jako czasu nauki **nie** traktowano okresu, jaki upłynął pomiędzy zakończeniem nauki w badanej szkole a ew. rozpoczęciem nauki w kolejnej (okresu *wakacji*).

Przy obliczaniu wskaźników nie były brane pod uwagę prace świadczone bez umowy.

- **praca_nauka_0m, praca_nauka_1m, praca_nauka_2m, praca_nauka_3m, praca_nauka_4m, praca_nauka_5m, praca_nauka_6m, praca_nauka_7m, praca_nauka_8m, praca_nauka_9m** – zmienne tekstowe opisujące status edukacyjno-zawodowy absolwenta w kolejnych miesiącach, licząc od miesiąca ukończenia szkoły; tworzone na podstawie odpowiedzi na pytania ZP2A, ZP2B, ZP2C, ZP2F, ZP2G, SP6C1, SP6C2, SP6D, SP6E1, SP6E2, SP6H, PP6C1, PP6C2, PP6D, PP6F, PP6F2, PG2C, PG2D, PG2E, PG2F:
 - „tylko praca” – absolwent w danym miesiącu pracował (w dowolnej formie, wykluczając zatrudnienie bez umowy) i jednocześnie się nie uczył;
 - „praca i edukacja” – absolwent w danym miesiącu pracował (w dowolnej formie, wykluczając zatrudnienie bez umowy) i jednocześnie nie uczył;
 - „tylko edukacja” – absolwent w danym miesiącu nie pracował (w dowolnej formie, wykluczając zatrudnienie bez umowy) i jednocześnie się uczył;
 - brak danych – wszystkie pozostałe sytuacje.

7.2.12. Wskaźniki opisujące bezrobocie i średnie wynagrodzenie w powiatach przygotowane na podstawie BDL GUS

- **powiat_bezrobocie_1m, powiat_bezrobocie_2m, powiat_bezrobocie_3m, powiat_bezrobocie_4m, powiat_bezrobocie_5m, powiat_bezrobocie_6m, powiat_bezrobocie_7m, powiat_bezrobocie_8m, powiat_bezrobocie_9m** – stopa

bezrobocia rejestrowanego w powiecie, w którym znajduje się szkoła (którą ukończył absolwent) w kolejnych miesiącach od miesiąca ukończenia szkoły:

- liczba od 0 do 100 [procent].
- **powiat_sr_wynagrodzenia_0r** – oszacowanie przeciętnego miesięcznego wynagrodzenia brutto w roku ukończenia szkoły w powiecie, w którym znajduje się szkoła (którą ukończył absolwent):
 - liczba [PLN].

7.3. Wskaźniki obliczane na poziomie zagregowanym

W odróżnieniu od wskaźników obliczanych na poziomie indywidualnym, wskaźniki obliczane na poziomie zagregowanym (z wyjątkiem tych opisanych w sekcji 7.3.1.) nie są typowymi zmiennymi liczbowymi lub tekstowymi, lecz są tworzone w formie obiektów o bardziej złożonej strukturze. Dla każdego ze wskaźników opisany został poniżej zakres przechowywanych w nim informacji i sposób ich obliczania. Techniczne szczegóły implementacji odpowiednich struktur danych opisane są w dokumentacji pakietu *MLASZdane* i nie będą tutaj rozważane.

7.3.1. Ogólne informacje o szkole/grupie porównawczej

- **SKZ_kod** – identyfikator szkoły, jednoznacznie identyfikujący placówkę.
 - Nie dotyczy wskaźników obliczonych dla *grup porównawczych*.
- **SKZ_typ** – typ szkoły zawodowej: „Szkoła policealna”, „Technikum” lub „Zasadnicza Szkoła Zawodowa”.
- **GRUPA_kod** – identyfikator *grupy porównawczej*.
- **liczba_zbadanych** – liczba absolwentów danej szkoły (lub *grupy porównawczej*), którzy znaleźli się w zbiorze z obliczonymi wskaźnikami na poziomie indywidualnym.
- **liczba_zbadanych_kobiet** – liczba kobiet wśród absolwentów danej szkoły (lub *grupy porównawczej*), którzy znaleźli się w zbiorze z obliczonymi wskaźnikami na poziomie indywidualnym.
- **liczba_szkol** – liczba szkół w *grupie porównawczej* lub 1 w zbiorze wskaźników szkół.
- **zawody** – ciąg znaków z (oddzielonymi przecinkami) nazwami zawodów (wartościami wskaźnika **UCZ_zawod** z poziomu indywidualnego), w których uczyli się w szkole absolwenci danej szkoły (lub *grupy porównawczej*), którzy znaleźli się w zbiorze z obliczonymi wskaźnikami na poziomie indywidualnym; zawody ułożone są w kolejności od najczęściej do najrzadziej występującego.

7.3.2. Wskaźniki statusu edukacyjno-zawodowego absolwentów

- **praca_nauka_0m, praca_nauka_1m, praca_nauka_2m, praca_nauka_3m, praca_nauka_4m, praca_nauka_5m, praca_nauka_6m, praca_nauka_7m, praca_nauka_8m, praca_nauka_9m** – wskaźniki opisujące liczbę absolwentów należących do danej grupy, o danym statusie edukacyjno-zawodowym, w kolejnych miesiącach od miesiąca ukończenia szkoły; zawierają następujące informacje:
 - **`n`** – liczba wszystkich absolwentów w grupie;
 - **`tylko pracują`** – liczba absolwentów w grupie, dla których wskaźniki indywidualne odpowiednio **praca_nauka_0m, praca_nauka_1m, praca_nauka_2m, praca_nauka_3m, praca_nauka_4m, praca_nauka_5m, praca_nauka_6m, praca_nauka_7m, praca_nauka_8m, praca_nauka_9m** przyjmują wartość „tylko praca”;
 - **`pracują i uczą się`** – liczba absolwentów w grupie, dla których wskaźniki indywidualne odpowiednio **praca_nauka_0m, praca_nauka_1m, praca_nauka_2m, praca_nauka_3m, praca_nauka_4m, praca_nauka_5m, praca_nauka_6m, praca_nauka_7m, praca_nauka_8m, praca_nauka_9m** przyjmują wartość „praca i edukacja”;
 - **`tylko się uczą`** – liczba absolwentów w grupie, dla których wskaźniki indywidualne odpowiednio **praca_nauka_0m, praca_nauka_1m, praca_nauka_2m, praca_nauka_3m, praca_nauka_4m, praca_nauka_5m, praca_nauka_6m, praca_nauka_7m, praca_nauka_8m, praca_nauka_9m** przyjmują wartość „tylko edukacja”;
 - **`nie pracują i nie uczą się`** – liczba absolwentów w grupie, dla których wskaźniki indywidualnych odpowiednio **praca_nauka_0m, praca_nauka_1m, praca_nauka_2m, praca_nauka_3m, praca_nauka_4m, praca_nauka_5m, praca_nauka_6m, praca_nauka_7m, praca_nauka_8m, praca_nauka_9m** są brakami danych;

7.3.3. Wskaźniki zdawalności egzaminów

- **egz_zaw_zdawalnosc** – udział absolwentów, którzy w roku ukończenia szkoły uzyskali dyplom potwierdzający kwalifikacje w zawodzie, którego uczyli się w szkole wśród absolwentów danej grupy:
 - **`n`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **egz_zaw_zdany** nie jest brakiem danych;
 - **`zdawalność`** - średnia wartość wskaźnika indywidualnego **egz_zaw_zdany** w grupie (liczba z przedziału [0;1]).
- **matura_zdawalnosc** udział absolwentów, którzy w roku ukończenia szkoły zdali maturę wśród absolwentów danej grupy:

- **`n`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **matura_zdany** nie jest brakiem danych (w przypadku szkół innych niż technika zawsze 0);
- **`zdawalność`** - średnia wartość wskaźnika indywidualnego **egz_zaw_zdany** w grupie (liczba z przedziału [0;1], a w przypadku szkół innych niż technika brak danych).

7.3.4. Wskaźniki opisujące warunki zatrudnienia absolwentów

- **praca_przed_ukonczeniem_szkoły** – liczba absolwentów w grupie, którzy pracowali przed ukończeniem (badanej) szkoły:
 - **`n`** – liczba wszystkich absolwentów w grupie;
 - **`praca przed ukończeniem szkoły`** - liczba absolwentów w grupie, dla których wskaźnik indywidualny **praca_przed_ukonczeniem_szkoły_pierwsza** przyjmuje wartość PRAWDA.
- **praca_czas_rozp** – parametry rozkładu wskaźnika indywidualnego **praca_czas_rozp_pierwsza**, opisującego liczbę miesięcy pomiędzy ukończeniem szkoły a podjęciem pierwszej pracy (przez absolwentów, którzy nie pracowali w momencie kończenia szkoły, ale później ją podjęli):
 - **`n`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **praca_czas_rozp_pierwsza** nie jest brakiem danych;
 - **`średnia`** – średnia wartość wskaźnika indywidualnego **praca_czas_rozp_pierwsza** w grupie;
 - **`mediana`** – mediana wartości wskaźnika indywidualnego **praca_czas_rozp_pierwsza** w grupie;
 - **`1.kwartyl`** – 1. kwartyl wartości wskaźnika indywidualnego **praca_czas_rozp_pierwsza** w grupie;
 - **`3.kwartyl`** – 3. kwartyl wartości wskaźnika indywidualnego **praca_czas_rozp_pierwsza** w grupie.
- **praca_forma_pierwsza, praca_forma_ostatnia** – wskaźniki opisujące liczbę absolwentów w grupie pracujących w poszczególnych formach zatrudnienia odpowiednio w pierwszej pracy wykonywanej po ukończeniu szkoły i w ostatniej pracy wykonywanej w okresie dziewięciu miesięcy od miesiąca ukończenia szkoły; zawierają następujące informacje:
 - **`n`** – liczba absolwentów w grupie, którzy pracowali po ukończeniu szkoły (tj. tych, dla których wskaźniki indywidualne odpowiednio **pg2gh.1_pierwsza** lub **pg2gh.1_ostatnia** nie są brakami danych);

- **`umowa o pracę na czas nieokreślony`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pg2gh.2_pierwsza** lub **pg2gh.2_ostatnia** przyjmuje wartość PRAWDA;
 - **`umowa o pracę na czas określony`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pg2gh.1_pierwsza** lub **pg2gh.1_ostatnia** przyjmuje wartość PRAWDA;
 - **`umowa zlecenia/o dzieło`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pg2gh.4_pierwsza** lub **pg2gh.4_ostatnia** przyjmuje wartość PRAWDA;
 - **`samozatrudnienie`** – liczba absolwentów w grupie, dla których wskaźniki indywidualne odpowiednio (**pg2gh.5_pierwsza** lub **pg2gh.6_pierwsza**) lub (**pg2gh.5_ostatnia** lub **pg2gh.6_ostatnia**) przyjmują wartości PRAWDA;
 - **`rolnicy indywidualni`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pg2gh.7_pierwsza** lub **pg2gh.7_ostatnia** przyjmuje wartość PRAWDA;
 - **`przez agencję pracy tymczasowej`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pg2gh.3_pierwsza** lub **pg2gh.3_ostatnia** przyjmuje wartość PRAWDA;
 - **`staż lub praktyka`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pg2gh.8_pierwsza** lub **pg2gh.8_ostatnia** przyjmuje wartość PRAWDA;
 - **`inna (lub nieznana)`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pg2gh.9_pierwsza** lub **pg2gh.9_ostatnia** przyjmuje wartość PRAWDA.
- **praca_forma2_pierwsza** – wskaźnik analogiczny do **praca_forma_pierwsza**, ale wyróżniająco mniejszą liczbę form zatrudnienia:
- **`n`** – liczba absolwentów w grupie, którzy pracowali po ukończeniu szkoły (tj. tych, dla których wskaźnik indywidualny odpowiednio **pg2gh.1_pierwsza** nie jest brakiem danych);
 - **`umowa o pracę na czas nieokreślony`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **pg2gh.2_pierwsza** przyjmuje wartość PRAWDA;
 - **`umowa o pracę na czas określony`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **pg2gh.1_pierwsza** przyjmuje wartość PRAWDA;
 - **`umowa zlecenia/o dzieło`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **pg2gh.4_pierwsza** przyjmuje wartość PRAWDA;

- **`inna (lub nieznana)`** – liczba absolwentów w grupie, dla których przynajmniej jeden ze wskaźników indywidualnych **pg2gh.3_pierwsza**, **pg2gh.5_pierwsza**, **pg2gh.6_pierwsza**, **pg2gh.7_pierwsza**, **pg2gh.8_pierwsza**, **pg2gh.9_pierwsza** przyjmuje wartość PRAWDA.

- **praca_forma_6m**, **praca_forma_9m**, **praca_forma2_6m**, **praca_forma2_9m** – wskaźniki opisują liczbę absolwentów w grupie pracujących w poszczególnych formach zatrudnienia odpowiednio w szóstym i w dziewiątym miesiącu od miesiąca ukończenia szkoły; przygotowywane analogicznie, jak wskaźniki odnoszące się do pierwszej i ostatniej pracy, z tym że obliczane na podstawie wskaźników indywidualnych odpowiednio (**pg2gh.1_6m**, **pg2gh.2_6m**, **pg2gh.3_6m**, **pg2gh.4_6m**, **pg2gh.5_6m**, **pg2gh.6_6m**, **pg2gh.7_6m**, **pg2gh.8_6m**, **pg2gh.9_6m**) i (**pg2gh.1_9m**, **pg2gh.2_9m**, **pg2gh.3_9m**, **pg2gh.4_9m**, **pg2gh.5_9m**, **pg2gh.6_9m**, **pg2gh.7_9m**, **pg2gh.8_9m**, **pg2gh.9_9m**).

- **praca_forma2_bu_6m**, **praca_forma2_bu_9m** – wskaźniki opisują liczbę absolwentów w grupie, którzy odpowiednio w szóstym i w dziewiątym miesiącu od miesiąca ukończenia szkoły **nie uczyli się** i pracowali w poszczególnych formach zatrudnienia; przygotowywane analogicznie, jak wskaźniki **praca_forma2_6m**, **praca_forma2_9m**, ale na zawężonej podgrupie absolwentów:
 - **`n`** – liczba absolwentów w grupie, dla których wskaźniki indywidualne odpowiednio **praca_nauka_6m** lub **praca_nauka_9m** nie przyjmują wartości „tylko praca” oraz wskaźniki indywidualne odpowiedni **pg2gh.1_6m** lub **pg2gh.1_9m** nie są brakami danych;
 - pozostałe informacje analogicznie, jak we wcześniej opisanych wskaźnikach odnoszących się do formy zatrudnienia, ale z wykluczeniem uczących się na podstawie wskaźników indywidualnych odpowiednio **praca_nauka_6m** lub **praca_nauka_9m**.

- **praca_zamieszkanie_pierwsza**, **praca_zamieszkanie_ostatnia** – wskaźniki opisują liczbę absolwentów w badanej grupie, którzy wykonywali pracę w Polsce lub za granicą, odpowiednio w odniesieniu do pierwszej pracy wykonywanej po ukończeniu (badanej) szkoły lub ostatniej pracy wykonywanej w okresie dziewięciu miesięcy od miesiąca ukończenia szkoły; zawierają następujące informacje:
 - **`n`** – liczba absolwentów w grupie, którzy pracowali po ukończeniu szkoły (tj. tych, dla których wskaźniki indywidualne odpowiednio **pg2i.1_pierwsza** lub **pg2i.1_ostatnia** nie są brakami danych);
 - **`w Polsce`** – liczba absolwentów w grupie, dla których wskaźniki indywidualne odpowiednio (**pg2i.1_pierwsza** lub **pg2i.2_pierwsza**) lub (**pg2i.1_ostatnia** lub **pg2i.2_ostatnia**) przyjmują wartości PRAWDA;
 - **`za granicą`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pg2i.3_pierwsza** lub **pg2i.3_ostatnia** przyjmuje wartość PRAWDA;

- **`nieznane`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pg2i.9_pierwsza** lub **pg2i.9_ostatnia** przyjmuje wartość PRAWDA.

- **praca_zamieszkanie_6m, praca_zamieszkanie_9m** – wskaźniki opisują liczbę absolwentów w grupie liczbę absolwentów w badanej grupie, którzy wykonywali pracę w Polsce lub za granicą odpowiednio w szóstym i w dziewiątym miesiącu od miesiąca ukończenia szkoły; przygotowywane analogicznie, jak wskaźniki odnoszące się do pierwszej i ostatniej pracy, z tym że obliczane na podstawie wskaźników indywidualnych odpowiednio (**pg2i.1_6m, pg2i.2_6m, pg2i.3_6m, pg2i.9_6m**) i (**pg2i.1_9m, pg2i.2_9m, pg2i.3_9m, pg2i.9_9m**).

- **praca_zgodna_z_wykształceniem_pierwsza, praca_zgodna_z_wykształceniem_ostatnia** – wskaźniki opisujące liczbę absolwentów w grupie, którzy w dany sposób ocenili zgodność swojej pracy z zawodem, którego uczyli się w badanej szkole, odpowiednio w odniesieniu do pierwszej pracy wykonywanej po ukończeniu (badanej) szkoły lub ostatniej pracy wykonywanej w okresie dziewięciu miesięcy od miesiąca ukończenia szkoły; zawierają następujące informacje:
 - **`n`** – liczba absolwentów w grupie, którzy dokonali oceny zgodności wykonywanej pracy z wykształceniem (tj. tych, dla których wskaźniki indywidualne odpowiednio **pio1_pierwsza** lub **pio1_ostatnia** nie są brakami danych);
 - **`zgodna`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pio1_pierwsza** lub **pio1_ostatnia** przyjmuje wartość 1;
 - **`niezgodna, ale wymaga podobnych kwalifikacji`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pio1_pierwsza** lub **pio1_ostatnia** przyjmuje wartość 2;
 - **`niezgodna, wymaga innych kwalifikacji`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pio1_pierwsza** lub **pio1_ostatnia** przyjmuje wartość 3;
 - **`praca, w której wykształcenie nie ma znaczenia`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pio1_pierwsza** lub **pio1_ostatnia** przyjmuje wartość 4.

- **praca_zarobki_pierwsza, praca_zarobki_ostatnia** – wskaźniki opisują rozkład zarobków netto absolwentów w grupie odpowiednio w odniesieniu do pierwszej pracy wykonywanej po ukończeniu (badanej) szkoły lub ostatniej pracy wykonywanej w okresie dziewięciu miesięcy od miesiąca ukończenia szkoły; zawierają następujące informacje:
 - **`n`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny odpowiednio **pio4_pierwsza** lub **pio4_ostatnia** nie jest brakiem danych;
 - **`średnia`** – średnia wartość wskaźnika indywidualnego odpowiednio **pio4_pierwsza** lub **pio4_ostatnia** w grupie;

- **`mediana`** – mediana wartości wskaźnika indywidualnego odpowiednio **pio4_pierwsza** lub **pio4_ostatnia** w grupie;
 - **`1.kwartyl`** – 1. kwartyl wartości wskaźnika indywidualnego odpowiednio **pio4_pierwsza** lub **pio4_ostatnia** w grupie;
 - **`3.kwartyl`** – 3. kwartyl wartości wskaźnika indywidualnego odpowiednio **pio4_pierwsza** lub **pio4_ostatnia** w grupie;
 - **`tlo`** - średnia wartość w grupie wskaźnika indywidualnego **powiat_sr_wynagrodzenia_0r**, opisującego przeciętne miesięczne wynagrodzenie brutto w powiecie, w którym znajduje się szkoła ukończona przez absolwenta, przemnożona przez 0,71 (przyjęte jako przybliżenie przejścia z zarobków brutto na zarobki netto).
- **praca_spełnienie_oczekiwan_ostatnia** opisuje liczbę absolwentów w grupie, którzy w dany sposób ocenili spełnianie ich oczekiwań odnośnie do pracy przez ostatnią pracę wykonywaną przez nich w okresie dziewięciu miesięcy od miesiąca ukończenia szkoły:
- **`n`** – liczba absolwentów w grupie, którzy dokonali oceny spełnienia ich oczekiwań przez wykonywaną pracę (tj. tych, dla których wskaźnik indywidualny **po5_ostatnia** nie jest brakiem danych);
 - **`niedostateczny`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **po5_ostatnia** przyjmuje wartość 1;
 - **`dopuszczający`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **po5_ostatnia** przyjmuje wartość 2;
 - **`dostateczny`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **po5_ostatnia** przyjmuje wartość 3;
 - **`dobry`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **po5_ostatnia** przyjmuje wartość 4;
 - **`bardzo dobry`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **po5_ostatnia** przyjmuje wartość 5;
 - **`celujący`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **po5_ostatnia** przyjmuje wartość 6.

7.3.5. Wskaźniki opisujące długość okresów pracy i bezrobocia

- **praca_czas_p9m_rozklad** – opisuje liczbę absolwentów, którzy przepracowali daną liczbę miesięcy w okresie dziewięciu miesięcy od miesiąca ukończenia szkoły:
- **`n`** – liczba absolwentów w grupie;
 - **`0 miesięcy`** - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_p9m** zawiera się w przedziale [0; 1/9);

- `1 miesiąc` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_p9m** zawiera się w przedziale [1/9; 2/9);
 - `2 miesiące` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_p9m** zawiera się w przedziale [2/9; 3/9);
 - `3 miesiące` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_p9m** zawiera się w przedziale [3/9; 4/9);
 - `4 miesiące` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_p9m** zawiera się w przedziale [4/9; 5/9);
 - `5 miesięcy` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_p9m** zawiera się w przedziale [5/9; 6/9);
 - `6 miesięcy` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_p9m** zawiera się w przedziale [6/9; 7/9);
 - `7 miesięcy` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_p9m** zawiera się w przedziale [7/9; 8/9);
 - `8 miesięcy` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_p9m** zawiera się w przedziale [8/9; 1);
 - `9 miesięcy` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_p9m** wynosi 1.
- **praca_czas_uop_p9m_rozklad** – opisuje liczbę absolwentów, którzy przepracowali daną liczbę miesięcy w okresie dziewięciu miesięcy od miesiąca ukończenia szkoły pracując w formie umowy o pracę:
- informacje o analogicznej strukturze, jak we wcześniej opisanym wskaźniku **praca_czas_p9m_rozklad**, ale obliczane na podstawie wartości wskaźnika indywidualnego **praca_czas_uop_p9m**.
- **praca_czas_gdy_bez_nauki_p9m_rozklad** – opisuje liczbę absolwentów, dla których w określonych przedziałach zawierają się wartości wskaźnika indywidualnego **praca_czas_gdy_bez_nauki_p9m**, opisującego udział czasu pracy nie łączonej z nauką w stosunku do łącznego czasu, w którym absolwent się nie uczył, w ramach okresu dziewięciu miesięcy od miesiąca ukończenia szkoły:
- `n` – liczba absolwentów w grupie, dla których wskaźnik indywidualny **praca_czas_gdy_bez_nauki_p9m** nie jest brakiem danych;
 - `0% okresu` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_gdy_bez_nauki_p9m** zawiera się w przedziale [0; 0,2);

- `20% okresu` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_gdy_bez_nauki_p9m** zawiera się w przedziale [0,2; 0,4);
 - `40% okresu` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_gdy_bez_nauki_p9m** zawiera się w przedziale [0,4; 0,6);
 - `60% okresu` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_gdy_bez_nauki_p9m** zawiera się w przedziale [0,6; 0,8);
 - `80% okresu` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_gdy_bez_nauki_p9m** zawiera się w przedziale [0,8; 1);
 - `100% okresu` - liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego **praca_czas_gdy_bez_nauki_p9m** wynosi 1.
- **praca_czas_gdy_bez_nauki_uop_p9m_rozklad** – opisuje liczbę absolwentów, którzy przez daną liczbę miesięcy w okresie dziewięciu miesięcy od miesiąca ukończenia szkoły byli zarejestrowanymi bezrobotnymi:
 - informacje o analogicznej strukturze, jak we wcześniej opisanym wskaźniku **praca_czas_gdy_bez_nauki_p9m_rozklad**, ale obliczane na podstawie wartości wskaźnika indywidualnego **praca_czas_gdy_bez_nauki_uop_p9m**.
- **bezrobocie_czas_p9m_rozklad** – opisuje liczbę absolwentów, dla których w określonych przedziałach zawierają się wartości wskaźnika indywidualnego **bezrobocie_czas_p9m**, opisującego udział czasu bycia zarejestrowanym bezrobotnym i jednocześnie nie ucznia się w stosunku do łącznego czasu, w którym absolwent się nie uczył, w ramach okresu dziewięciu miesięcy od miesiąca ukończenia szkoły:
 - informacje o analogicznej strukturze, jak we wcześniej opisanym wskaźniku **praca_czas_p9m_rozklad**, ale obliczane na podstawie wartości wskaźnika indywidualnego **bezrobocie_czas_p9m_rozklad**.
- **bezrobocie_czas_gdy_bez_nauki_p9m_rozklad** – opisuje liczbę absolwentów, dla których w określonych przedziałach zawierają się wartości wskaźnika indywidualnego **bezrobocie_czas_gdy_bez_nauki_p9m**, opisującego udział czasu bycia zarejestrowanym bezrobotnym i jednocześnie nie ucznia się w stosunku do łącznego czasu, w którym absolwent się nie uczył, w ramach okresu dziewięciu miesięcy od miesiąca ukończenia szkoły:
 - informacje o analogicznej strukturze, jak we wcześniej opisanym wskaźniku **praca_czas_gdy_bez_nauki_p9m_rozklad**, ale obliczane na podstawie wartości wskaźnika indywidualnego **bezrobocie_czas_gdy_bez_nauki_p9m**.

- **praca_czas_p9m, praca_czas_gdy_bez_nauki_p9m, praca_czas_uop_p9m, praca_czas_gdy_bez_nauki_uop_p9m, bezrobocie_czas_p9m, bezrobocie_czas_gdy_bez_nauki_p9m** – wskaźniki opisują te same cechy grupy absolwentów, co opisane powyżej wskaźniki o analogicznych nazwach z dodanym przyrostkiem „_rozklad”, przy czym zamiast zdyskretyzowanych rozkładów liczebności zawierają informacje o wartościach parametrów poziomu wartości odpowiednich wskaźników indywidualnych (o tych samych nazwach co przygotowywane wskaźniki zagregowane); zawierają następujące informacje:
 - **`n`** – liczba absolwentów w grupie, dla których odpowiedni wskaźnik indywidualny nie jest brakiem danych;
 - **`średnia`** – średnia wartość odpowiedniego wskaźnika indywidualnego w grupie;
 - **`mediana`** – mediana wartości odpowiedniego wskaźnika indywidualnego w grupie;
 - **`1.kwartyl`** – 1. kwartyl wartości odpowiedniego wskaźnika indywidualnego w grupie;
 - **`3.kwartyl`** – 3. kwartyl wartości odpowiedniego wskaźnika indywidualnego w grupie.

- **praca_czas_gdy_nauka_p9m, praca_czas_gdy_nauka_uop_p9m, bezrobocie_czas_gdy_nauka_p9m** – wskaźniki opisujące rozkłady udziału czasu odpowiednio: pracy (w dowolnej formie oprócz pracy bez umowy) łączonej z nauką, pracy w formie umowy o pracę łączonej z nauką, bycia zarejestrowanym bezrobotnym łączonego z nauką w czasie nauki w okresie dziewięciu miesięcy od miesiąca ukończenia szkoły; tworzone na podstawie wskaźników indywidualnych o takich samych nazwach:
 - informacje o analogicznej strukturze, jak we wcześniej opisanym wskaźniku **praca_czas_p9m**, ale obliczane na podstawie wartości wskaźników indywidualnych odpowiednio **praca_czas_gdy_nauka_p9m**, **praca_czas_gdy_nauka_uop_p9m** i **bezrobocie_czas_gdy_nauka_p9m**.

7.3.6. Wskaźniki opisujące bezrobocie absolwentów

- **bezrobocie_1m, bezrobocie_2m, bezrobocie_3m, bezrobocie_4m, bezrobocie_5m, bezrobocie_6m, bezrobocie_7m, bezrobocie_8m, bezrobocie_9m** – wskaźniki opisujące liczbę absolwentów w grupie, którzy w kolejnych miesiącach od miesiąca ukończenia szkoły byli zarejestrowanymi bezrobotnymi; zawierają następujące informacje:
 - **`n`** – liczba absolwentów w grupie, którzy w danym miesiącu uczyli się, pracowali, lub byli zarejestrowanymi bezrobotnymi (tj. liczba absolwentów w grupie, dla których wskaźniki indywidualne odpowiednio **bezrobocie_1m**, **bezrobocie_2m**, **bezrobocie_3m**, **bezrobocie_4m**, **bezrobocie_5m**, **bezrobocie_6m**, **bezrobocie_7m**, **bezrobocie_8m** lub **bezrobocie_9m** nie są brakami danych);

- **`zarejestrowani bezrobotni`** – liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego odpowiednio **bezrobocie_1m**, **bezrobocie_2m**, **bezrobocie_3m**, **bezrobocie_4m**, **bezrobocie_5m**, **bezrobocie_6m**, **bezrobocie_7m**, **bezrobocie_8m** lub **bezrobocie_9m** jest równa 1;
- **`pracujący lub kontynuujący naukę`** – liczba absolwentów w grupie, dla których wartość wskaźnika indywidualnego odpowiednio **bezrobocie_1m**, **bezrobocie_2m**, **bezrobocie_3m**, **bezrobocie_4m**, **bezrobocie_5m**, **bezrobocie_6m**, **bezrobocie_7m**, **bezrobocie_8m** lub **bezrobocie_9m** jest równa 0.

7.3.7. Wskaźniki opisujące kontynuację nauki przez absolwentów

- **nauka_6m, nauka_9m** – wskaźniki opisują liczbę absolwentów w grupie, którzy odpowiednio w szóstym lub dziewiątym miesiącu licząc od miesiąca ukończenia szkoły uczyli się w szkołach poszczególnych typów; zawierają następujące informacje:
 - **`n`** – liczba absolwentów w grupie, którzy w odpowiednio szóstym lub dziewiątym miesiącu licząc od miesiąca ukończenia szkoły uczyli się w szkole jednego z niżej wymienionych typów (tj. absolwentów, dla których wskaźniki indywidualne odpowiednio **nauka_6m** lub **nauka_9m** nie są brakami danych);
 - **`studia stacjonarne`** – liczba absolwentów w grupie, dla których wskaźniki indywidualne odpowiednio **nauka_6m** lub **nauka_9m** przyjmują wartość „studia stacjonarne”;
 - **`studia niestacjonarne`** – liczba absolwentów w grupie, dla których wskaźniki indywidualne odpowiednio **nauka_6m** lub **nauka_9m** przyjmują wartość „studia niestacjonarne”;
 - **`szkoła policealna`** – liczba absolwentów w grupie, dla których wskaźniki indywidualne odpowiednio **nauka_6m** lub **nauka_9m** przyjmują wartość „szkoła policealna”;
 - **`LO dla dorosłych`** – liczba absolwentów w grupie, dla których wskaźniki indywidualne odpowiednio **nauka_6m** lub **nauka_9m** przyjmują wartość „LO dla dorosłych”.
- **studia_gdzie_pierwsze** – opisuje uczelnie i kierunki studiów wyższych najczęściej wybierane przez absolwentów grupy jako miejsca kontynuacji kształcenia po ukończeniu szkoły; zawierają następujące informacje:
 - **`n`** – liczba absolwentów w grupie, którzy po ukończeniu szkoły rozpoczęli naukę na studiach (tj. wskaźnik indywidualny **studia_uczelnia_pierwsze** nie jest dla nich brakiem danych); dla szkół typów innych niż technikum zawsze 0;
 - **`najczęściej wybierane kierunki`** – ciąg znaków z (oddzielonymi przecinkami) nazwami kierunków (wartościami wskaźnika

studia_kierunek_pierwsze z poziomu indywidualnego), na których rozpoczęli naukę absolwenci danej grupy; nazwy kierunków ułożone są w kolejności od najczęściej do najrzadziej występującego wśród absolwentów danej grupy; nazwy kierunków, które wymieniło mniej niż 20% absolwentów danej grupy są pomijane;

- **`najczęściej wybierane uczelnie`** – ciąg znaków z (oddzielonymi przecinkami) nazwami uczelni (wartościami wskaźnika **studia_uczelnia_pierwsze** z poziomu indywidualnego), na których rozpoczęli naukę absolwenci danej grupy; nazwy uczelni ułożone są w kolejności od najczęściej do najrzadziej występującej wśród absolwentów danej grupy; nazwy uczelni, które wymieniło mniej niż 20% absolwentów danej grupy są pomijane.
- **studia_odplatnosc_pierwsze** – opisuje liczbę absolwentów grupy, którzy po ukończeniu szkoły podjęli studia płatne i studia bezpłatne:
 - **`n`** – liczba absolwentów w grupie, którzy po ukończeniu szkoły rozpoczęli naukę na studiach i udzielili odpowiedzi na pytanie o to, czy były to studia odpłatne (tj. wskaźnik indywidualny **studia_bezplatne_pierwsze** nie jest dla nich brakiem danych); dla szkół typów innych niż technikum zawsze 0;
 - **`studiujący na studiach bezpłatnych`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **studia_bezplatne_pierwsze** przyjmuje wartość 0;
 - **`studiujący na studiach płatnych`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **studia_bezplatne_pierwsze** przyjmuje wartość 1.
- **studia_tryb_pierwsze** – opisuje liczbę absolwentów grupy, którzy po ukończeniu szkoły podjęli studia w poszczególnych trybach:
 - **`n`** – liczba absolwentów w grupie, którzy po ukończeniu szkoły rozpoczęli naukę na studiach i udzielili odpowiedzi na pytanie o tryb odbywania studiów (tj. wskaźnik indywidualny **studia_tryb_pierwsze** nie jest dla nich brakiem danych); dla szkół typów innych niż technikum zawsze 0;
 - **`studiujący na studiach dziennych`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **studia_tryb_pierwsze** przyjmuje wartość 1;
 - **`studiujący na studiach wieczorowych`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **studia_tryb_pierwsze** przyjmuje wartość 2;
 - **`studiujący na studiach zaocznych`** – liczba absolwentów w grupie, dla których wskaźnik indywidualny **studia_tryb_pierwsze** przyjmuje wartość 3.

8. Charakterystyka respondentów

Poniżej znajdują się rozkłady podstawowych cech respondentów ze względu na typ szkoły. Rozkłady te pochodzą ze wszystkich modułów ilościowych zastosowanych w pierwszej rundzie monitoringu:

- Uczniowie zbadani metodą ankiety audytoryjnej (Tabele: 8.1. – 8.4),
- Uczniowie zbadani metodą wywiadu CAPI w szkole (Tabele: 8.5 – 8.9),
- Absolwenci zbadani metodą wywiadu CAPI w II fali badania (Tabele: 8.10 – 8.14),
- Dyrektorzy szkół zbadani metodą wywiadu CAPI w szkole (Tabele: 8.15 – 8.18).

Wszystkie poniższe dane uwzględniają wagi poststratyfikacyjne do analizy w podziałach ze względu na typ szkoły.

Tabela 8.1. Rozkład płci uczniów zbadanych metodą ankiety audytoryjnej

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
	liczba	procent	liczba	procent	liczba	procent
Kobieta	2270	38,0%	1531	42,2%	2056	78,2%
Mężczyzna	3703	62,0%	2098	57,8%	572	21,8%
Brak odpowiedzi	3	0,0%	0	0,0%	0	0,0%
Ogółem	5976	100,0%	3629	100,0%	2628	100,0%

Tabela 8.2. Rozkład roku urodzenia uczniów zbadanych metodą ankiety audytoryjnej

Rok	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
	liczba	procent	liczba	procent	liczba	procent
1987	1	0,0%	0	0,0%	17	0,6%
1988	0	0,0%	0	0,0%	65	2,5%
1989	0	0,0%	0	0,0%	81	3,1%
1990	0	0,0%	0	0,0%	95	3,6%
1991	0	0,0%	1	0,0%	133	5,1%
1992	11	0,2%	0	0,0%	247	9,4%
1993	1	0,0%	0	0,0%	335	12,8%
1994	8	0,1%	2	0,0%	316	12,0%
1995	16	0,3%	23	0,6%	493	18,8%
1996	260	4,4%	159	4,4%	660	25,1%
1997	1419	23,7%	3400	93,7%	141	5,4%
1998	4166	69,7%	35	1,0%	11	0,4%
1999	23	0,4%	0	0,0%	5	0,2%
Odmowa odpowiedzi	70	1,2%	9	0,2%	29	1,1%
Ogółem	5976	100,0%	3629	100,0%	2628	100,0%

Tabela 8.3. Rozkład obszaru kształcenia, w jakim kształcą się uczniowie zbadani metodą ankiety audytoryjnej

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Administracyjno-usługowy	1630	27,3%	915	25,2%	664	25,3%
Budowlany	665	11,1%	521	14,3%	30	1,1%
Elektryczno-elektroniczny	511	8,5%	421	11,6%	99	3,8%
Mechaniczny i górniczo-hutniczy	1428	23,9%	609	16,8%	6	0,2%
Rolniczo-leśny z ochroną środowiska	86	1,4%	492	13,6%	119	4,5%
Artystyczny	1	0,0%	5	0,1%	5	0,2%
Turystyczno-gastronomiczny	1579	26,4%	660	18,2%	87	3,3%
Medyczno-społeczny	0	0,0%	0	0,0%	1565	59,6%
Brak danych	77	1,3%	7	0,2%	52	2,0%
Ogółem	5976	100,0%	3629	100,0%	2628	100,0%

Tabela 8.4. Rozkład województw, w jakich znajduje się szkoła, w której kształcą się uczniowie zbadani metodą ankiety audytoryjnej

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Dolnośląskie	97	1,6%	68	1,9%	41	1,5%
Kujawsko-pomorskie	537	9,0%	347	9,5%	111	4,2%
Lubelskie	286	4,8%	123	3,4%	517	19,7%
Lubuskie	362	6,1%	30	0,8%	7	0,3%
Łódzkie	146	2,4%	187	5,1%	128	4,9%
Małopolskie	809	13,5%	200	5,5%	228	8,7%
Mazowieckie	629	10,5%	702	19,3%	275	10,5%
Opolskie	72	1,2%	25	0,7%	64	2,4%
Podkarpackie	257	4,3%	116	3,2%	88	3,3%
Podlaskie	64	1,1%	163	4,5%	424	16,1%
Pomorskie	723	12,1%	278	7,7%	205	7,8%
Śląskie	402	6,7%	207	5,7%	159	6,1%
Świętokrzyskie	350	5,9%	461	12,7%	82	3,1%
Warmińsko-mazurskie	597	10,0%	81	2,2%	115	4,4%
Wielkopolskie	424	7,1%	517	14,3%	152	5,8%
Zachodniopomorskie	221	3,7%	125	3,4%	32	1,2%
Ogółem	5976	100,0%	3629	100,0%	2628	100,0%

Tabela 8.5. Rozkład płci uczniów zbadanych metodą wywiadu CAPI

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Kobieta	1577	41,3%	1847	37,4%	1004	82,0%
Mężczyzna	2241	58,7%	3096	62,6%	220	18,0%
Ogółem	3818	100,0%	4943	100,0%	1224	100,0%

Tabela 8.6. Rozkład roku urodzenia uczniów zbadanych metodą wywiadu CAPI

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
1977	0	0,0%	1	0,0%	0	0,0%
1987	1	0,0%	6	0,1%	2	0,1%
1988	0	0,0%	0	0,0%	22	1,8%
1989	0	0,0%	0	0,0%	33	2,7%
1990	0	0,0%	1	0,0%	37	3,1%
1991	0	0,0%	0	0,0%	71	5,8%
1992	0	0,0%	0	0,0%	168	13,7%
1993	0	0,0%	0	0,0%	138	11,2%
1994	2	0,0%	5	0,1%	122	10,0%
1995	12	0,3%	17	0,3%	189	15,4%
1996	148	3,9%	193	3,9%	336	27,5%
1997	805	21,1%	4695	95,0%	80	6,5%
1998	2825	74,0%	20	0,4%	6	0,5%
1999	23	0,6%	6	0,1%	1	0,1%
Brak danych	2	0,1%	0	0,0%	20	1,6%
Ogółem	3818	100,0%	4943	100,0%	1224	100,0%

Tabela 8.7. Rozkład obszaru kształcenia, w jakim kształcą się uczniowie zbadani metodą wywiadu CAPI

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Administracyjno-usługowy	1220	32,0%	1205	24,4%	216	17,6%
Budowlany	475	12,4%	659	13,3%	13	1,0%
Elektryczno-elektroniczny	342	8,9%	1006	20,3%	27	2,2%
Mechaniczny i górniczo-hutniczy	820	21,5%	810	16,4%	1	0,1%
Rolniczo-leśny z ochroną środowiska	75	2,0%	268	5,4%	61	5,0%
Artystyczny	0	0,0%	3	0,1%	0	0,0%
Turystyczno-gastronomiczny	883	23,1%	988	20,0%	47	3,9%
Medyczno-społeczny	0	0,0%	3	0,1%	839	68,5%
Brak danych	2	0,1%	2	0,0%	21	1,7%
Ogółem	3818	100,0%	4943	100,0%	1224	100,0%

Tabela 8.8. Rozkład województw, w jakich znajduje się szkoła, w której kształcą się uczniowie zbadani metodą wywiadu CAPI

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Dolnośląskie	34	0,9%	178	3,6%	16	1,3%
Kujawsko-pomorskie	518	13,6%	475	9,6%	20	1,6%
Lubelskie	188	4,9%	157	3,2%	493	40,3%
Lubuskie	119	3,1%	262	5,3%	3	0,2%
Łódzkie	61	1,6%	287	5,8%	13	1,1%
Małopolskie	1023	26,8%	470	9,5%	162	13,2%
Mazowieckie	231	6,0%	457	9,3%	41	3,3%
Opolskie	95	2,5%	132	2,7%	63	5,1%
Podkarpackie	259	6,8%	301	6,1%	3	0,2%
Podlaskie	102	2,7%	100	2,0%	257	21,0%
Pomorskie	127	3,3%	290	5,9%	33	2,7%
Śląskie	370	9,7%	625	12,6%	44	3,6%
Świętokrzyskie	55	1,4%	347	7,0%	0	0,0%
Warmińsko-mazurskie	124	3,3%	105	2,1%	42	3,4%
Wielkopolskie	454	11,9%	612	12,4%	35	2,8%
Zachodniopomorskie	57	1,5%	145	2,9%	0	0,0%
Ogółem	3818	100,0%	4943	100,0%	1224	100,0%

Tabela 8.9. Rozkład wielkości miejscowości zamieszkania uczniów uczniowie zbadanych metodą wywiadu CAPI

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Wieś nie będąca siedzibą gminy	1674	43,8%	1838	37,2%	326	26,6%
Wieś będąca siedzibą gminy	877	23,0%	733	14,8%	125	10,2%
Małe miasto (do 20 tys. mieszkańców)	441	11,6%	663	13,4%	140	11,5%
Średnie miasto (od 21 do 100 tys. mieszkańców)	509	13,3%	960	19,4%	208	17,0%
Duże miasto (od 101 do 500 tys. mieszkańców)	243	6,4%	508	10,3%	307	25,1%
Bardzo duże miasto (powyżej 500 tys. mieszkańców)	73	1,9%	241	4,9%	118	9,7%
Ogółem	3818	100,0%	4943	100,0%	1224	100,0%

Tabela 8.10. Rozkład płci absolwentów zbadanych metodą wywiadu CAPI w II fali badania

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Kobieta	1218	43,2%	1378	35,4%	741	80,4%
Mężczyzna	1598	56,8%	2518	64,6%	180	19,6%
Ogółem	2816	100,0%	3896	100,0%	921	100,0%

Tabela 8.11. Rozkład roku urodzenia absolwentów zbadanych metodą wywiadu CAPI w II fali badania

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
1977	0	0,0%	1	0,0%	0	0,0%
1987	1	0,0%	6	0,2%	2	0,2%
1988	0	0,0%	0	0,0%	21	2,3%
1989	0	0,0%	0	0,0%	30	3,3%
1990	0	0,0%	0	0,0%	28	3,0%
1991	0	0,0%	0	0,0%	32	3,5%
1992	0	0,0%	0	0,0%	157	17,0%
1993	0	0,0%	0	0,0%	92	10,0%
1994	2	0,1%	3	0,1%	89	9,7%
1995	4	0,1%	11	0,3%	159	17,2%
1996	102	3,6%	139	3,6%	229	24,8%
1997	555	19,7%	3716	95,4%	66	7,1%
1998	2130	75,6%	16	0,4%	3	0,3%
1999	19	0,7%	4	0,1%	0	0,0%
Brak danych	3	0,1%	0	0,0%	13	1,4%
Ogółem	2816	100,0%	3896	100,0%	921	100,0%

Tabela 8.12. Rozkład obszaru kształcenia, w jakim kształcili się absolwentów zbadani metodą wywiadu CAPI w II fali badania

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Administracyjno-usługowy	921	32,7%	961	24,7%	162	17,6%
Budowlany	390	13,8%	537	13,8%	12	1,4%
Elektryczno-elektroniczny	171	6,1%	819	21,0%	20	2,2%
Mechaniczny i górniczo-hutniczy	602	21,4%	668	17,1%	1	0,1%
Rolniczo-leśny z ochroną środowiska	57	2,0%	208	5,3%	60	6,5%
Artystyczny	0	0,0%	3	0,1%	0	0,0%
Turystyczno-gastronomiczny	672	23,9%	697	17,9%	28	3,0%
Medyczno-społeczny	0	0,0%	2	0,0%	624	67,8%
Brak danych	2	0,1%	1	0,0%	14	1,5%
Ogółem	2816	100,0%	3896	100,0%	921	100,0%

Tabela 8.13. Rozkład województw, w jakich znajduje się szkoła, w której kształcili się absolwenci zbadani metodą wywiadu CAPI w II fali badania

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Dolnośląskie	17	0,6%	167	4,3%	15	1,6%
Kujawsko-pomorskie	281	10,0%	307	7,9%	9	0,9%
Lubelskie	135	4,8%	156	4,0%	373	40,4%
Lubuskie	114	4,0%	221	5,7%	3	0,4%
Łódzkie	52	1,9%	261	6,7%	11	1,2%
Małopolskie	784	27,9%	440	11,3%	145	15,7%
Mazowieckie	227	8,1%	395	10,1%	39	4,2%
Opolskie	79	2,8%	71	1,8%	40	4,3%
Podkarpackie	281	10,0%	280	7,2%	3	0,3%
Podlaskie	79	2,8%	49	1,3%	175	19,0%
Pomorskie	62	2,2%	193	5,0%	34	3,7%
Śląskie	244	8,7%	480	12,3%	32	3,4%
Świętokrzyskie	45	1,6%	263	6,8%	0	0,0%
Warmińsko-mazurskie	27	1,0%	36	0,9%	25	2,8%
Wielkopolskie	365	12,9%	523	13,4%	19	2,1%
Zachodniopomorskie	23	0,8%	54	1,4%	0	0,0%
Ogółem	2816	100,0%	3896	100,0%	921	100,0%

Tabela 8.14. Rozkład wielkości miejscowości zamieszkania absolwentów zbadanych metodą wywiadu CAPI w II fali badania

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Wieś nie będąca siedzibą gminy	1259	44,7%	1482	38,0%	260	28,2%
Wieś będąca siedzibą gminy	666	23,7%	568	14,6%	89	9,7%
Małe miasto (do 20 tys. mieszkańców)	339	12,0%	521	13,4%	100	10,8%
Średnie miasto (od 21 do 100 tys. mieszkańców)	321	11,4%	739	19,0%	176	19,1%
Duże miasto (od 101 do 500 tys. mieszkańców)	183	6,5%	369	9,5%	218	23,7%
Bardzo duże miasto (powyżej 500 tys. mieszkańców)	48	1,7%	216	5,6%	78	8,5%
Ogółem	2816	100,0%	3896	100,0%	921	100,0%

Tabela 8.15. Rozkład płci respondentów zbadanych na podstawie kwestionariusza do wywiadu CAPI z dyrektorem szkoły

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Kobieta	340	67,3%	300	63,9%	300	75,7%
Mężczyzna	165	32,7%	170	36,1%	96	24,3%
Ogółem	505	100,0%	470	100,0%	396	100,0%

Tabela 8.16. Rozkład stażu pracy w danej szkole wśród respondentów zbadanych na podstawie kwestionariusza do wywiadu CAPI z dyrektorem szkoły

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
5 lat lub mniej	52	10,2%	46	9,8%	113	28,6%
6-10 lat	57	11,3%	44	9,4%	84	21,2%
11-20 lat	161	31,8%	148	31,6%	103	26,0%
21-30 lat	159	31,4%	145	30,9%	64	16,1%
31 lat lub więcej	61	12,1%	76	16,1%	15	3,8%
Brak danych/nie pamiętam	16	3,1%	10	2,2%	17	4,2%
Ogółem	505	100,0%	470	100,0%	396	100,0%

Tabela 8.17. Rozkład stażu pracy jako dyrektor danej szkoły wśród respondentów zbadanych na podstawie kwestionariusza do wywiadu CAPI z dyrektorem szkoły

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
5 lat lub mniej	172	34,1%	166	35,4%	176	44,5%
6-10 lat	110	21,8%	96	20,4%	82	20,6%
11-20 lat	113	22,4%	103	21,9%	82	20,7%
21-30 lat	18	3,6%	22	4,6%	6	1,4%
31 lat lub więcej	2	0,4%	3	0,5%	1	0,3%
Respondent nie jest dyrektorem*	84	16,6%	78	16,5%	39	9,9%
Brak danych/nie pamiętam	6	1,2%	3	0,6%	10	2,6%
Ogółem	505	100,0%	470	100,0%	396	100,0%

*Procedura badania pozwala na to, by wywiad był przeprowadzony z inną osobą niż dyrektor szkoły, w przypadku gdy dyrektor delegował to zadanie na innego pracownika szkoły (np. wicedyrektora).

Tabela 8.18. Rozkład odpowiedzi na pytanie „czy szkoła należy do zespołu szkół” wśród respondentów zbadanych na podstawie kwestionariusza do wywiadu CAPI z dyrektorem szkoły

	Zasadnicza szkoła zawodowa		Technikum		Szkoła policealna	
Tak	454	90,0%	435	92,5%	158	39,9%
Nie	51	10,0%	35	7,5%	238	60,1%
Ogółem	505	100,0%	470	100,0%	396	100,0%

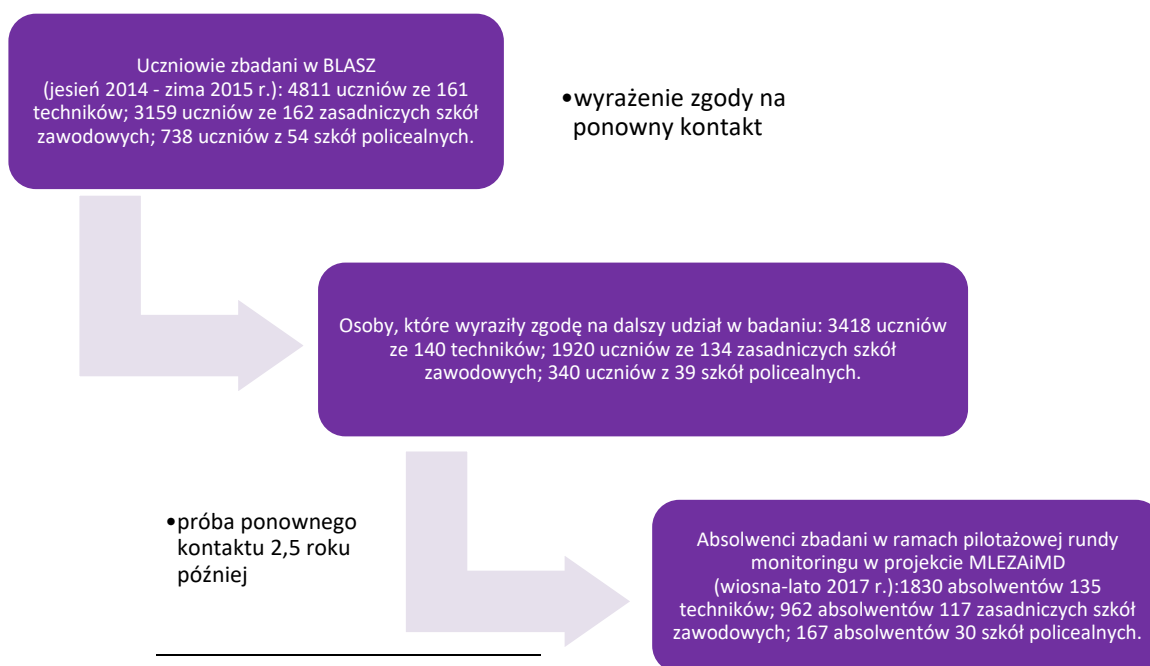
9. Runda pilotażowa

Opisywana we wcześniejszych rozdziałach pierwsza runda monitoringu została poprzedzona rundą pilotażową zrealizowaną na bazie wcześniejszego projektu. Poniżej przedstawione są główne założenia i efekty realizacji tej rundy.

9.1. Rola Badania Losów Absolwentów Szkół Zawodowych (BLASZ) w projekcie MLEZAiMD

Podstawą do realizacji pilotażowej rundy monitoringu losów absolwentów szkół zawodowych w ramach projektu MLEZAiMD była próba badawcza, na której w 2015 r. zrealizowano *Badanie Losów Absolwentów Szkół Zawodowych (BLASZ)*⁴. Próba badania BLASZ objęła osoby, które wiosną 2015 r. były uczniami ostatnich klas polskich techników i zasadniczych szkół zawodowych lub szkół policealnych (tych ostatnich bez względu na klasę/semestr). Z udziału w badaniu wykluczono osoby starsze niż 29 lat. Z badania wyłączono również szkoły artystyczne, tj. kształcące w zawodach przypisanych do artystycznego obszaru kształcenia. Realizacja wywiadów w szkołach w ramach badania BLASZ trwała od listopada 2014 r. do lutego 2015 r.

Rysunek 9.1. Relacje pomiędzy próbą BLASZ a próbą zrealizowaną w pilotażowej rundzie monitoringu w ramach projektu MLEZAiMD



⁴ Badanie było przeprowadzone przez IBE w ramach projektu systemowego Badanie jakości i efektywności edukacji oraz instytucjonalizacja zaplecza badawczego współfinansowanego ze środków Europejskiego Funduszu Społecznego w ramach Programu Operacyjnego Kapitał Ludzki, Priorytet III: Wysoka jakość systemu oświaty, Poddziałanie 3.1.1 Tworzenie warunków i narzędzi do monitorowania, ewaluacji i badań systemu oświaty.

9.2. Narzędzia

W każdej wylosowanej szkole przeprowadzano wywiad z dyrektorem szkoły na podstawie standaryzowanego kwestionariusza (treść kwestionariusza znaleźć można w Aneksie 1a). W badaniu uczniów wykorzystywano dwa narzędzia. Nominalnie wszyscy badani wypełniali krótką ankietę audytoryjną (treść ankiety znaleźć można w Aneksie 1b), w której zadawano pytania o uczestnictwo w praktykach zawodowych, plany związane z przystąpieniem do egzaminów zawodowych oraz ocenę różnych aspektów jakości kształcenia w szkole respondentów oraz o wykształcenie ich rodziców. W przybliżeniu 70% osób wypełniających ankietę audytoryjną zostało dodatkowo zbadanych przy pomocy znacznie obszerniejszego kwestionariusza (Aneks 1c), metodą wspomaganego komputerowo wywiadu osobistego z ankietą (tzw. CAPI). Łącznie zbadano 8708 osób z 377 szkół. Jednakże tylko część z nich zgodziła się wziąć udział w badaniu ich dalszych losów edukacyjnych i zawodowych. O ile w technikach zgody takiej udzieliło 71% badanych uczniów, o tyle w ZSZ niespełna 61%, a w szkołach policealnych tylko nieco ponad 46%. W związku z tym w ramach projektu MLEZAiMD możliwe było podjęcie próby kontaktu z 5678 osobami z 313 szkół.

Właśnie z tymi osobami podjęto próbę przeprowadzenia wywiadu, mającego na celu określenie ich losów edukacyjnych i zawodowych po ukończeniu szkoły zawodowej, w ramach pilotażowej rundy monitoringu realizowanej w ramach projektu MLEZAiMD. Badanie prowadzone było w okresie od 13 maja do 30 września 2017 r., a więc w około 2,5 roku po realizacji BLASZ, co dla większości uczestników odpowiadało okresowi 2 lat od zakończenia nauki w szkole zawodowej, w której byli uczniami w trakcie realizacji projektu BLASZ. Za realizację wywiadów z absolwentami odpowiadała agencja badawcza PBS Sp. z o.o., w ramach oddzielnego projektu *Monitorowanie losów absolwentów szkół zawodowych - pierwsza runda* (POWR.02.15.00-IP.02-00-003/16). Wywiady (kwestionariusz znaleźć można w Aneksie 1d) udało się przeprowadzić z 2959 osobami, co stanowi nieco ponad 52% tych, którzy wyrazili zgodę na dalszy udział i 34% spośród wszystkich uczestników BLASZ.

Tabela 9.1. Liczba osób zbadanych w BLASZ i pilotażowej rundzie monitoringu w ramach projektu MLEZAiMD

	Zbadani w 2014/2015 r. w BLASZ	Wyrazili zgodę na dalszy udział w badaniu	Zbadani w 2017 r. w MLEZAiMD
Technika	4811	3418	1830
w stosunku do liczby zbadanych w BLASZ	100,0%	71,0%	38,0%
w stosunku do liczby wyrażających zgodę		100,0%	53,5%
Zasadnicze szkoły zawodowe	3159	1920	962
w stosunku do liczby zbadanych w BLASZ	100,0%	60,8%	30,5%
w stosunku do liczby wyrażających zgodę		100,0%	50,1%

Szkoły Policealne	738	340	167
w stosunku do liczby zbadanych w BLASZ	100,0%	46,1%	22,6%
w stosunku do liczby wyrażających zgodę		100,0%	49,1%
Łącznie	8708	5678	2959
w stosunku do liczby zbadanych w BLASZ	100,0%	65,2%	34,0%
w stosunku do liczby wyrażających zgodę		100,0%	52,1%

Źródło: obliczenia własne na podstawie danych sondażowych BLASZ i MLEZAiMD

9.3. Przyczyny uniemożliwiające uzyskanie wywiadów

W odniesieniu do realizacji badania absolwentów, przeprowadzonego w ramach projektu MLEZAiMD w 2017 r. możliwe jest podanie bardziej szczegółowych informacji na temat przyczyn, dla których nie udało się uzyskać wywiadów, jako że były one odnotowywane w dokumentacji dotyczącej realizacji terenowej badania, realizowanej przez firmę PBS. Najważniejszą przeszkodą w realizacji badania okazały się niestety odmowy udzielenia wywiadu przez respondentów. Odnosząc liczbę odmów do liczby wszystkich osób, które dwa i pół roku wcześniej zgodziły się na dalszy udział w badaniu, wystąpiły one w przypadku aż 18% absolwentów techników i zasadniczych szkół zawodowych oraz 14% absolwentów szkół policealnych. Jeśli jednak wziąć pod uwagę tylko tych uczestników BLASZ, z którymi ankieterom udało się nawiązać osobisty kontakt, okazuje się, że spotykali się oni z odmową aż w jednej czwartej przypadków.

Poważne trudności sprawiały również niemożliwość zastania jakiegokolwiek osoby pod podanym adresem oraz zmiany miejsca zamieszkania badanych w czasie, jaki upłynął od momentu, gdy podawali oni swoje dane kontaktowe. W przypadku absolwentów zasadniczych szkół zawodowych zaznaczył się też problem większej liczby błędnie podanych adresów kontaktowych, uniemożliwiających identyfikację miejsca zamieszkania. Około 4% osób, z którymi starano się skontaktować, przez cały okres badania przebywała poza granicami kraju. Zasadne wydaje się przyjęcie założenia, że większość z nich stanowili emigranci, choć nie jesteśmy niestety w stanie stwierdzić, kiedy, gdzie i na jak długo zdecydowali się oni wyjechać.

Warto też odnotować, że negatywny wpływ na liczbę zrealizowanych wywiadów miał fakt, że beneficjent projektu obejmującego realizację terenową badania, nie podjął próby kontaktu z niemal 6% osób, których dane kontaktowe otrzymał.

Tabela 9.2. Ostateczny status prób realizacji wywiadów z pilotażowej rundy monitoringu z wyszczególnieniem przyczyn niezrealizowania wywiadów

	Technika		Zasadnicze szkoły zawodowe		Szkoły policealne	
Wywiady zrealizowane	1830	53,5%	962	50,1%	167	49,1%
Wywiady niezrealizowane:						
respondent przebywa za granicą	142	4,2%	86	4,5%	13	3,8%
respondent odmówił udzielenia wywiadu	618	18,1%	339	17,7%	48	14,1%
niemożliwość ustalenia obecnego adresu respondenta	197	5,8%	120	6,3%	28	8,2%
pod danym adresem nie udało się z nikim skontaktować	298	8,7%	164	8,5%	46	13,5%
błędny adres lub adres niedostępny	46	1,3%	79	4,1%	5	1,5%
wykonawca badania nie podjął próby kontaktu z respondentem	215	6,3%	96	5,0%	18	5,3%
inne powody niezrealizowania	72	2,1%	74	3,9%	15	4,4%
Suma	3418	100,0%	1920	100,0%	340	100,0%

Źródło: obliczenia własne na podstawie danych sondażowych MLEZAiMD

9.4. Ograniczenia próby zrealizowanej w Badaniu Losów Absolwentów Szkół Zawodowych

Realizacja *Badania Losów Absolwentów Szkół Zawodowych* natrafiła na liczne trudności, które niestety w negatywny sposób wpłynęły na reprezentatywność próby i możliwość uogólniania wyników prowadzonych na niej analiz. Najważniejszy z nich to bardzo duże problemy z uzyskaniem zgody na realizację badania w szkołach policealnych. Blisko dwie trzecie pierwotnie wylosowanych do badania szkół tego typu odmówiło wzięcia w nim udziału. W konsekwencji zrealizowana próba szkół policealnych jest nieliczna i wyraźnie różni się od populacji uczniów tego typu szkół w zakresie częstości występowania poszczególnych obszarów kształcenia zawodowego. Niedoreprezentowany jest obszar administracyjno-usługowy, a bardzo nadreprezentowany jest obszar turystyczno-gastronomiczny. Drugi poważny problem, dotyczący szkół wszystkich typów polega na tym, że podczas realizacji badania na terenie województwa kujawsko-pomorskiego w ogóle nie zebrano zgód na dalszy udział uczestników w badaniu. W konsekwencji, próba objęta pilotażową rundą monitoringu

w ogóle nie obejmuje absolwentów szkół z województwa kujawsko-pomorskiego. Należy też odnotować, że w wyniku nieprzestrzegania założonych procedur przy doborze respondentów w ramach szkół, do próby trafiły również nieliczne osoby starsze niż 29 lat (dotyczyło to niemal wyłącznie szkół policealnych).

9.5. Ograniczenia próby zrealizowanej w pilotażowej rundzie monitoringu

Na opisane powyżej ograniczenia nakłada się dodatkowo kwestia dużej terytorialnej nierównomierności, jeśli chodzi o skuteczność w uzyskiwaniu wywiadów podczas realizacji drugiej fali pilotażowej rundy monitoringu (realizowanej w ramach projektu konkursowego). O ile w województwie świętokrzyskim udało się zrealizować badanie z 73% osób, a w mazowieckim z nieco ponad 70% osób, które w BLASZ wyraziły zgodę na ponowny kontakt, o tyle w województwie pomorskim było to już tylko nieco ponad 25%, a w zachodniopomorskim niecałe 16% takich osób.

W przypadku szkół policealnych najpoważniejszy problem stanowi jednak bardzo niewielka liczebność zrealizowanej próby: 167 osób z zaledwie 30 różnych placówek. Choć skuteczność w uzyskiwaniu wywiadów była w przypadku absolwentów szkół policealnych tylko niewiele niższa niż dla absolwentów szkół pozostałych typów (49% wobec 50% dla ZSZ i 54% dla techników). Jednak wobec niewielkiej liczby zgód na dalszy udział w badaniu, jakie udało się dwa lata wcześniej zebrać w ramach BLASZ od uczniów szkół policealnych, ostatecznie zrealizowana próba jest bardzo nieliczna. W związku z niewielką liczbą badanych w praktyce niemożliwe jest prowadzenie w ramach pilotażowej rundy monitoringu bardziej zaawansowanych analiz, odnoszących się do uczniów szkół policealnych.

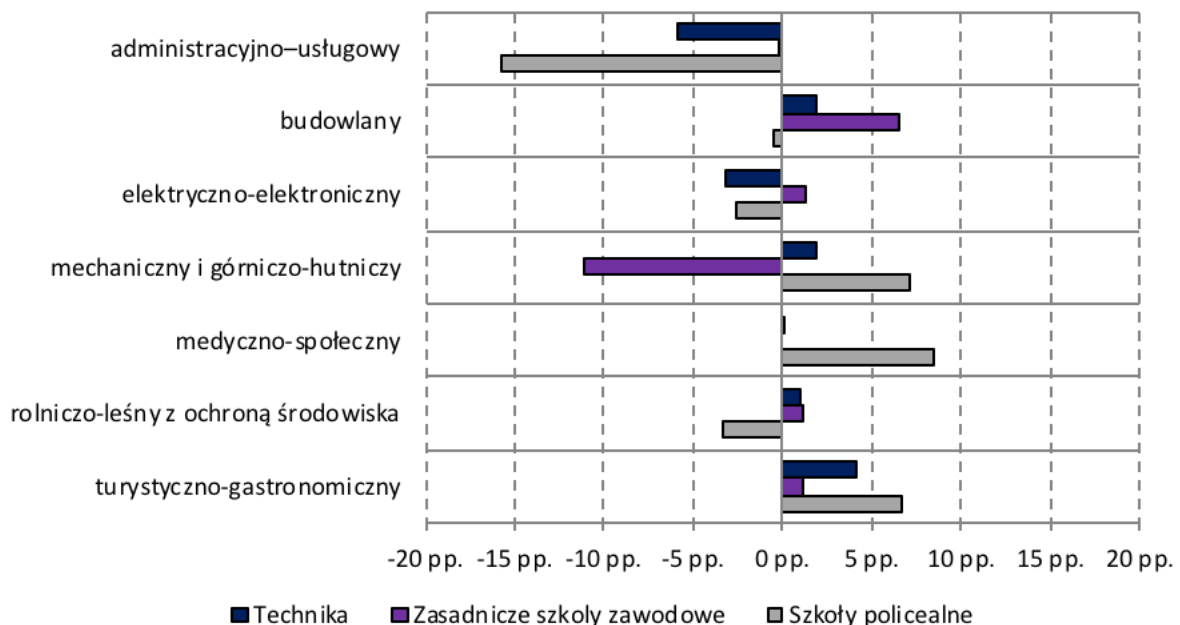
9.6. Możliwość uogólniania wniosków z analizy wyników pilotażowej rundy monitoringu

Z opisanych powyżej powodów próba absolwentów zrealizowana w badaniu MLEZAiMD nie może być traktowana jako reprezentatywna dla populacji absolwentów uczniów ostatnich klas polskich szkół zawodowych w roku szkolnym 2014/2015. W wyniku kumulacji oddziaływania różnych czynników mających wpływ na możliwość uzyskania wywiadu z poszczególnymi uczniami i absolwentami, jej kompozycja różni się wyraźnie od składu populacji, w szczególności ze względu na udział osób kształconych w poszczególnych obszarach kształcenia zawodowego oraz reprezentację poszczególnych regionów kraju. Nie można więc na jej podstawie formułować wniosków nt. ogólnej sytuacji w ramach populacji, z której pierwotnie dobierana była próba (BLASZ), np. odnośnie odsetka kontynuujących naukę pracujących, czy poziomu zarobków. Niemniej cały czas pozostaje ona cennym źródłem informacji pozwalającym badać uwarunkowania zróżnicowania dalszych ścieżek edukacyjnych i zawodowych absolwentów, pozwalając określić najważniejsze czynniki ryzyka i sukcesu dla przebiegu ich karier.

Analizując skład zrealizowanej próby MLEZAiMD warto zwrócić uwagę na trzy charakterystyki, których rozkłady w populacji uczniów ostatnich klas szkół zawodowych w roku szkolnym 2014/2015 są znane (dają się obliczyć na podstawie danych z Systemu Informacji Oświatowej): obszar kształcenia zawodowego, do jakiego przypisany jest

nauczany zawód, lokalizację szkoły, ze względu na region kraju oraz płeć. Na wykresie po**Błąd! Nie można odnaleźć źródła odwołania.** przedstawione zostały różnice pomiędzy procentowym udziałem osób uczących się w poszczególnych typach szkół (w roku szkolnym 2014/2015) w ramach poszczególnych obszarów kształcenia, a analogicznymi udziałami procentowymi w zrealizowanej próbie MLEZAI^{MD}. Widać, że wśród absolwentów techników wyraźnie niedoreprezentowane (o 6 pkt. proc.) są osoby wykształcone w obszarze administracyjno-usługowym, a w nieco mniejszym stopniu (o 3 pkt. proc.) w obszarze elektryczno-elektronicznym. Nadreprezentowany jest natomiast obszar turystyczno-gastronomiczny (o 4 pkt. proc.) oraz, choć tu różnice są już niewielkie i nie przekraczają 2 pkt. proc.: budowlany, mechaniczny i górnictwo-hutnictwo oraz rolniczy z ochroną środowiska. W przypadku absolwentów zasadniczych szkół zawodowych zdecydowanie niedoreprezentowane są z kolei absolwenci kierunków z obszaru mechanicznego i górnictwo-hutnictwa (aż o 11 pkt. proc.). Nadreprezentowany jest za to obszar budowlany (o 7 pkt. proc.). W pozostałych obszarach odchylenia od rozkładu populacyjnego są już niewielkie. Największe różnice pomiędzy zrealizowaną próbą a populacją odnotowywane są w przypadku absolwentów szkół policealnych, przy czym po części jest to konsekwencją jej niewielkiej liczebności. Należy odnotować tu bardzo duże niedoreprezentowanie obszaru administracyjno-usługowego (aż o 16 pkt. proc.), przy nadreprezentacji obszarów medyczno-społecznego (o 8 pkt. proc.), mechanicznego i górnictwo-hutnictwa oraz turystyczno-gastronomicznego (oba o około 7 pkt. proc.).

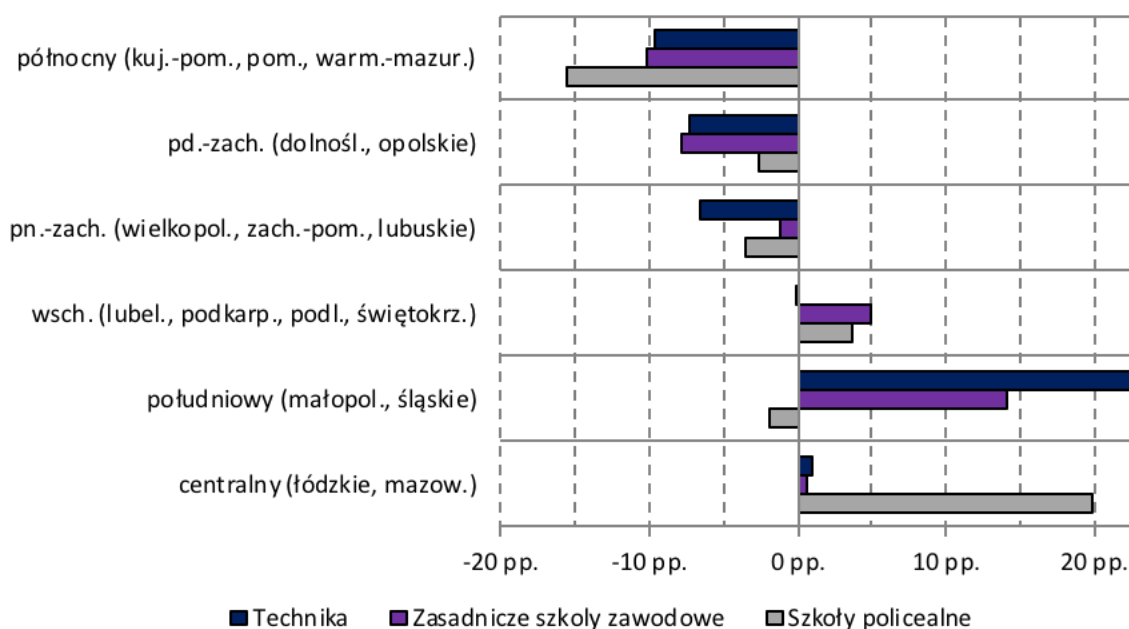
Rysunek 9.2. Reprezentacja obszarów kształcenia w zrealizowanej próbie MLEZAI^{MD} w porównaniu do populacji uczniów ostatnich klas szkół zawodowych w roku szkolnym 2014/2015 (z wyłączeniem obszaru artystycznego)



Źródło: obliczenia własne na podstawie danych sondażowych MLEZAI^{MD} oraz SIO. Obszar kształcenia jest dla szkół danego typu nadreprezentowany, jeśli słupek na wykresie skierowany jest w prawo (wartości dodatnie), a niedoreprezentowany, jeśli słupek skierowany jest w lewo (wartości ujemne).

Jeśli chodzi o terytorialne rozmieszczenie próby zrealizowanej w badaniu MLEZAIiMD (rozpatrywane z punktu widzenia lokalizacji szkoły, w której badani uczyli się w roku szkolnym 2014/2015), dla wszystkich trzech typów szkół wyraźnie niedoreprezentowane są północne i zachodnie regiony Polski. W przypadku absolwentów techników i zasadniczych szkół zawodowych silnie nadreprezentowany jest za to region południowy (tj. województwa małopolskie i śląskie) – w przypadku techników aż o 22 pkt. proc. W przypadku absolwentów szkół policealnych w podobnie dużym stopniu (o 20 pkt. proc.) nadreprezentowany był region centralny.

Rysunek 9.3. Reprezentacja regionów (NUTS 1) w zrealizowanej próbie MLEZAIiMD w porównaniu do populacji uczniów ostatnich klas szkół zawodowych w roku szkolnym 2014/2015



Źródło: obliczenia własne na podstawie danych sondażowych MLEZAMID oraz SIO. Region jest dla szkół danego typu szkół nadreprezentowany, jeśli słupek na wykresie skierowany jest w prawo (wartości dodatnie), a niedoreprezentowany, jeśli słupek skierowany jest w lewo (wartości ujemne).

Udział kobiet i mężczyzn w próbie jest zbliżony do populacyjnego w zbadanej grupie absolwentów techników (61% mężczyzn w próbie wobec 60% w populacji). Wśród absolwentów zasadniczych szkół zawodowych i szkół policealnych nieco niedoreprezentowani są jednak mężczyźni – dla obu typów szkół różnica wynosi około 7 pkt. proc., przy czym o ile wśród zbadanych absolwentów zasadniczych szkół zawodowych mężczyźni stanowią 61%, o tyle wśród absolwentów szkół policealnych jedynie 23%.

Rozkłady omówionych powyżej zmiennych, a także wieku i wielkości miejscowości zamieszkania absolwentów, w zrealizowanej próbie absolwentów przedstawione są poniżej.

Tabela 9.3. Liczba zbadanych absolwentów w poszczególnych obszarach kształcenia zawodowego (wg zawodu, jakiego uczyli się w szkole, jako uczniowie której zostali zbadani w BLASZ)

Obszar kształcenia zawodowego	Technika		Zasadnicze szkoły zawodowe		Szkoły policealne	
Administracyjno-usługowy	356	19,5%	256	26,6%	41	24,6%
Budowlany	196	10,7%	183	19,0%	0	0,0%
Elektryczno-elektroniczny	376	20,5%	69	7,2%	7	4,2%
Mechaniczny i górniczo-hutniczy	247	13,5%	211	21,9%	13	7,8%
Medyczno-społeczny	1	0,1%	0	0,0%	88	52,7%
Rolniczo-leśny z ochroną środowiska	165	9,0%	18	1,9%	4	2,4%
Turystyczno-gastronomiczny	486	26,6%	215	22,3%	14	8,4%
Nieznany	3	0,2%	10	1,0%	0	0,0%
Suma	1830	100,0%	962	100,0%	167	100,0%

Tabela 9.4. Liczba zbadanych absolwentów z poszczególnych regionów – wg lokalizacji szkoły jako uczniowie której zostali zbadani w BLASZ

Region (województwa)	Technika		Zasadnicze szkoły zawodowe		Szkoły policealne	
centralny (łódzkie, mazowieckie)	333	18,2%	141	14,7%	66	39,5%
południowy (małopolskie, śląskie)	810	44,3%	346	36,0%	32	19,2%
wschodni (lubelskie, podkarpackie, podlaskie, świętokrzyskie)	360	19,7%	199	20,7%	40	24,0%
północno-zachodni (wielkopolskie, zachodniopomorskie,	173	9,5%	181	18,8%	19	11,4%

lubuskie)						
południowo-zachodni (dolnośląskie, opolskie)	42	2,3%	16	1,7%	10	6,0%
północny (kujawsko- pomorskie, pomorskie, warmińsko-mazurskie)	112	6,1%	79	8,2%	0	0,0%
Suma	1830	100,0%	962	100,0%	167	100,0%

Tabela 9.5. Rozkład płci wśród zbadanych absolwentów

Płeć	Technika		Zasadnicze szkoły zawodowe		Szkoły policealne	
	mężczyźni	1112	60,8%	591	61,4%	38
kobiety	718	39,2%	371	38,6%	129	77,2%
Suma	1830	100,0%	962	100,0%	167	100,0%

Tabela 9.6. Rozkład wieku wśród zbadanych absolwentów

Wiek (rocznikowy w 2017 r.)	Technika		Zasadnicze szkoły zawodowe		Szkoły policealne	
	19-20 lat	11	0,6%	38	4,0%	0
21 lat	33	1,8%	700	72,8%	0	0,0%
22 lata	1716	93,8%	186	19,3%	4	2,4%
23-27 lat	59	3,2%	35	3,6%	105	62,9%
28-32 lat	1	0,1%	0	0,0%	29	17,4%
33 lata i więcej	2	0,1%	1	0,1%	27	16,2%
brak danych	8	0,4%	2	0,2%	2	1,2%
Suma	1830	100,0%	962	100,0%	167	100,0%

Tabela 9.7. Rozkład wielkości miejscowości zamieszkania (w 2017 r.) wśród zbadanych absolwentów

Wielkość miejscowości zamieszkania	Technika		Zasadnicze szkoły zawodowe		Szkoły policealne	
Wieś niebędąca siedzibą gminy	766	41,9%	464	48,2%	40	24,0%
Wieś będącą siedzibą gminy	212	11,6%	80	8,3%	15	9,0%
Małe miasto (do 20 tys. mieszkańców)	181	9,9%	119	12,4%	24	14,4%
Średnie miasto (od 21 do 100 tys. mieszkańców)	311	17,0%	146	15,2%	29	17,4%
Duże miasto (od 101 do 500 tys. mieszkańców)	249	13,6%	120	12,5%	34	20,4%
Bardzo duże miasto (powyżej 500 tys. mieszkańców)	111	6,1%	33	3,4%	25	15,0%
Suma	1830	100,0%	962	100,0%	167	100,0%

10. Badanie jakościowe

Runda pierwsza monitoringu zakładała również przeprowadzenie badań jakościowych na podstawie narzędzi przygotowanych przez zespół badawczy IBE. W ramach projektu „Monitorowanie Losów Absolwentów Szkół Zawodowych – pierwsza runda” beneficjent projektu konkursowego miał za zadanie przeprowadzić 30 zogniskowanych wywiadów grupowych (FGI) oraz 70 indywidualnych wywiadów pogłębionych (IDI). Respondentami w badaniu byli dyrektorzy, nauczyciele, doradcy zawodowi oraz uczniowie i absolwenci szkół zawodowych (ZSZ oraz techników), jak również pracodawcy współpracujący i niewspółpracujący ze szkołami zawodowymi oraz przedstawiciele samorządów.

10.1. Narzędzia

W założeniach tematyka badań miała obejmować następujące zagadnienia:

- ścieżki edukacyjno-zawodowe,
- ofertę edukacyjną szkół zawodowych,
- zasoby szkół zawodowych i organizacji PNZ,
- uczniów i ich rodziców,
- pracę w szkole,
- współpracę szkół z pracodawcami,
- system placówek kształcenia zawodowego w powiecie,
- zasady nauczania w szkołach i placówkach oświatowych w powiecie,
- lokalny rynek pracy i losy absolwentów na lokalnym rynku pracy,
- współpracę szkół kształcących zawodowo i pracodawców,
- przyszłość kształcenia zawodowego w powiecie.

Na potrzeby badania opracowano odrębne scenariusze indywidualnych wywiadów pogłębionych dla 7 kategorii rozmówców. Scenariusze posiadały części wspólne, jednak zakres wywiadu każdorazowo dopasowano do zakresu wiedzy danej kategorii respondentów. W Tabeli 10.1 przedstawiającej obszary tematyczne dla każdej z grup respondentów znaleźć można również odniesienia do aneksów, w których zawarta jest treść scenariuszy.

Tabela 10.1. Obszary tematyczne indywidualnych wywiadów pogłębionych ze względu na kategorię respondenta

Kategoria respondentów	Obszary tematyczne IDI	Numer aneksu zawierającego scenariusz
Uczniowie szkół zawodowych	ścieżki edukacyjno-zawodowo, ocena doświadczeń zawodowych, plany na przyszłość	Aneks 3b
Absolwenci szkół zawodowych, którzy kończyli edukację w roku szkolnym 2015/2016	ścieżki edukacyjno-zawodowo, ocena doświadczeń zawodowych, plany na przyszłość	Aneks 3c
Dyrektorzy szkół zawodowych	współpraca z rodzicami i uczniami, zasoby szkoły i organizacja pnz, rynek pracy dla absolwentów	Aneks 3d
Nauczyciele przedmiotów zawodowych, w tym nauczycieli praktycznej nauki zawodu	współpraca z rodzicami i uczniami, zasoby szkoły i organizacja pnz, rynek pracy dla absolwentów	Aneks 3e
Doradcy edukacyjno-zawodowi pracujący w szkołach zawodowych	praca w szkole i związane z nią wyzwania, zasoby szkoły i organizacja pnz, rynek pracy dla absolwentów	Aneks 3f
Pracodawcy	lokalny rynek pracy i pracownicy, absolwenci w firmie – rekrutacja, doświadczenia, ocena, współpraca szkół z pracodawcami i plany na przyszłość	Aneks 3g
Przedstawiciele jednostek samorządu terytorialnego na poziomie powiatu odpowiedzialnych za sprawy edukacji	system placówek kształcenia zawodowego, oferta kształcenia zawodowego, lokalny rynek pracy i miejsce absolwentów na rynku pracy, współpraca szkół z pracodawcami, przyszłość kształcenia zawodowego w powiecie	Aneks 3h

W przypadku zogniskowanych wywiadów grupowych posłużono się jednym rodzajem scenariusza wywiadu dla wszystkich grup uczestników (Aneks 3a). Zakres wywiadów obejmował:

- Winiety z historiami uczniów – rekonstrukcja losów zawodowych i edukacyjnych, budowanie wariantów alternatywnych, identyfikacja czynników kształtujących przebieg ścieżek kariery absolwentów (osie czasu),
- Analiza SWOT lokalnego rynku pracy,
- Projektowanie serwisu internetowego z informacjami o losach absolwentów szkół zawodowych.

10.2. Założenia doboru próby badawczej

Badania realizowano w dwóch regionach, które reprezentowały odmienne warunki na lokalnym rynku pracy, jak i w zakresie organizacji systemu kształcenia zawodowego. Region I obejmował miasto Wrocław, z niskim poziomem bezrobocia i skoncentrowaną terytorialnie siecią szkół kształcących zawodowo. Region II obejmował powiat szydlowiecki wraz z powiatami przyległymi (starachowicki, radomski, skarżyski, przysuski, konecki - 77 szkół w 42 zespołach⁵), który charakteryzuje się relatywnie trudną sytuacją na rynku pracy i rozproszoną siecią szkół zawodowych. W przypadku Regionu II rekrutację uczestników zogniskowanych wywiadów grupowych prowadzono dodatkowo również na obszarze Radomia. Liczebności jednostek badawczych w ramach poszczególnych regionów oraz grup respondentów przedstawia Tabela 10.2.

Tabela 10.2. Liczebność wywiadów w podziale na region i techniki realizacji

	RAZEM		REGION 1 WROCLAW		REGION 2 6 POWIATÓW	
	FGI	IDI	FGI	IDI	FGI	IDI
Dyrektorzy szkół zawodowych	4	4	2	2	2	2
Nauczyciele przedmiotów zawodowych	6	4	3	2	3	2
Uczniowie szkół zawodowych	4	20	2	10	2	10
Absolwenci szkół zawodowych	6	24	3	12	3	12
Pracodawcy	6	10	3	5	3	5
Doradcy edukacyjno-zawodowi pracujący w szkołach zawodowych	2	4	1	2	1	2
Przedstawiciele jednostek samorządu terytorialnego na poziomie powiatu	2	4	1	2	1	2
RAZEM	30	70	15	35	11	35

⁵ W przypadku FGI realizowanych w Regionie II założono, iż co najmniej czterech respondentów powinno być przedstawicielami powiatu starachowickiego lub skarżyskiego, oraz co najmniej czterech powinno reprezentować pozostałe powiaty.

Dodatkowe kryteria doboru respondentów zakładały:

- optymalną liczebność uczestników FGI na poziomie 10-12 osób, przy czym dopuszczalna była realizacja badania w mniejszych grupach,
- różnicowanie w doborze respondentów ze względu na branże i wielkość firmy (dotyczy grupy pracodawców),
- w przypadku FGI z pracodawcami: uczestnicy każdej grupy powinni być zróżnicowani ze względu na doświadczenie we współpracy ze szkołami uczącymi zawodu,
- IDI z pracodawcami: przynajmniej 2 IDI z pracodawcami, którzy nie współpracują ze szkołami, ale potencjalnie zatrudniają ich absolwentów,
- w badaniu przedstawicieli samorządu: FGI z radnymi powiatowymi – minimum połowa osób powinna być z powiatowych Komisji Edukacji lub Komisji zajmującej się gospodarką,
- w badaniu przedstawicieli samorządu: FGI - REGION 2: co najmniej czterech respondentów powinno być przedstawicielami powiatu starachowickiego lub skarżyskiego, oraz co najmniej czterech powinno reprezentować pozostałe powiaty (szydłowiecki, radomski, przysuski lub konecki).

10.3. Realizacja badania w terenie

Badanie poprzedzone było centralnym szkoleniem, które odbyło się w dniu 27.02.2017 w Warszawie. Szkolenie prowadzone było przez pracowników beneficjenta projektu konkursowego przy współudziale przedstawicieli Instytutu Badań Edukacyjnych. W szkoleniu wzięło udział 7 moderatorów.

Rekrutacja respondentów oraz realizacja badania jakościowego została przeprowadzona zgodnie z założeniami harmonogramu projektu. W związku z utrudnioną rekrutacją związaną z zakończeniem roku szkolnego proces ten był ciągły i rozłożony w czasie. Moderatorzy na bieżąco otrzymywali listę respondentów zrekrutowanych do badania. Ostatnie IDI realizowane były przez moderatorów 30.08.2017. Wraz z pracami terenowymi, prowadzone były także prace nad transkrypcjami z wywiadów. Wywiady prowadzone były na terenie szkół, placówek samorządowych, w profesjonalnych salach fokusowych oraz w punktach użyteczności publicznej.

Moderatorzy prowadzili spotkania zgodnie z opracowanymi przez zespół badawczy IBE i skonsultowanymi z przedstawicielami beneficjenta scenariuszami. W przypadku FGI do rejestracji przebiegu badania wykorzystywane były kamera oraz dyktafon, w przypadku IDI moderatorzy korzystali z dyktafonu.

10.4. Liczba zrealizowanych wywiadów

W badaniu zrealizowano całość opisanej próby badawczej – zgodnie z przyjętymi założeniami. Przeprowadzono łącznie 30 FGI oraz 70 IDI. W sumie przebadano 232 respondentów, z czego 162 respondentów brało udział w FGI, natomiast 70 respondentów uczestniczyło w IDI. Ze względu na dostępność respondentów w trakcie trwania badania, jak również ze względu na okoliczności losowe niezależne od wykonawcy, FGI przeprowadzone były w grupach liczących od 4 do 9 respondentów. Szczególne trudności realizacyjne sprawiały grupy uczniów i absolwentów, którzy mimo rekrutacji od 12 do 14 osób na grupę, stawiali się na spotkaniu w zdecydowanie mniejszej liczbie.

10.5. Czas trwania wywiadów

Średni czas trwania FGI wynosił 1 godzinę 48 minut, przy czym czas najkrótszej grupy, którą stanowili uczniowie w regionie 2 to 1 godzina 11 minut, natomiast czas trwania najdłuższego FGI z nauczycielami przedmiotów zawodowych w regionie 1 to 3 godziny. W przypadku wywiadów IDI średni czas wynosił 43 minuty, przy czym najkrótszy wywiad z pracodawcą z regionu 2 trwał 25 minut, natomiast najdłuższy z dyrektorem szkoły zawodowej z regionu 1 trwał 2 godziny i 8 minut. Z analizy czasów poszczególnych wywiadów wynika, że najkrótszymi wywiadami były te przeprowadzone z uczniami i absolwentami szkół zawodowych, którzy są dopiero na początku swojej ścieżki rozwoju zawodowego.

10.6. Trudności realizacyjne

W toku badania Wykonawca napotkał trudności związane z rekrutacją respondentów, szczególnie w przypadku zogniskowanych wywiadów grupowych. Założenia odnośnie do liczby uczestników tychże spotkań okazały się dalece zbyt optymistyczne. Mimo zakładanej liczby 8-12 uczestników, ich faktyczna liczba wahała się od 4 do 9 osób. Problemy te szczególnie uwidoczniły się w przypadku grup absolwentów i uczniów, które charakteryzowała najwyższa absencja uczestników – mimo potwierdzenia umówienia się na spotkanie.

Pojawiły się również trudności związane z dotarciem do przedstawicieli doradców edukacyjno-zawodowych, którzy pracowaliby w badanych szkołach. *Problemy z rekrutacją tej grupy wynikały głównie z trudności w operacjonalizacji pojęcia stanowiska przez dyrekcję szkoły. W wielu szkołach dyrektorzy deklarowali, że w ich placówce nie ma osoby pełniącej taką funkcję, co znacznie wydłużało proces dotarcia do respondentów, spełniających kryteria udziału w badaniu⁶.* Wykonawca starał się przezwyciężyć problemy rekrutacyjne wykorzystując do tych zadań najbardziej doświadczonych rekruterów, którzy byli odpowiedzialni za poszczególne regiony realizacji badania.

⁶ PBS (2018). *Raport z realizacji terenowej badania – moduł jakościowy „Monitorowanie Losów Absolwentów Szkół Zawodowych – pierwsza runda”* Sopot.

Literatura cytowana

Haman J. (2016), *Metodologia doboru próby uczniów ostatnich klas szkół zawodowych w badaniu Monitorowanie Losów Edukacyjno-Zawodowych Absolwentów i Młodych Dorosłych*, dokument wewnętrzny projektu.

Koniewski M. (2016), *Ekspertyza dotycząca metodologii badań sondażowych wykorzystywanych w badaniach losów absolwentów*, dokument wewnętrzny projektu.

Koniewski M. (2016), *Ekspertyza dotycząca wskaźników i metod analizy statystycznej wykorzystywanych w badaniach losów absolwentów*, dokument wewnętrzny projektu.

PBS (2018). *Raport z realizacji terenowej badania – moduł jakościowy „Monitorowanie Losów Absolwentów Szkół Zawodowych – pierwsza runda”* Sopot.

Aneks – narzędzia badawcze

Poniżej zaprezentowana została lista narzędzi badawczych, które wykorzystywane były w pilotażowej i pierwszej rundzie monitoringu. Stanowią one załączniki do niniejszego raportu.

- Aneks 1a – runda pilotażowa: BLASZ kwestionariusz do wywiadu z dyrektorem szkoły
- Aneks 1b – runda pilotażowa: BLASZ ankieta audytoryjna z uczniem
- Aneks 1c – runda pilotażowa: BLASZ kwestionariusz do wywiadu z uczniem
- Aneks 1d – runda pilotażowa: kwestionariusz do wywiadu z absolwentem
- Aneks 2a – runda pierwsza: kwestionariusz do wywiadu z dyrektorem szkoły
- Aneks 2b – runda pierwsza: ankieta audytoryjna z uczniem
- Aneks 2c – runda pierwsza: kwestionariusz do wywiadu z uczniem
- Aneks 2d – runda pierwsza: kwestionariusz do wywiadu z absolwentem
- Aneks 3a – Formularz zgody rodzica/opiekuna prawnego na udział dziecka w badaniu audytoryjnym oraz przetwarzanie danych osobowych
- Aneks 3b – Formularz zgody ucznia pełnoletniego na udział w badaniu audytoryjnym oraz przetwarzanie danych osobowych
- Aneks 3c – Formularz zgody rodzica/opiekuna prawnego na udział dziecka w badaniu panelowym oraz przetwarzanie danych osobowych

- Aneks 3d – Formularz zgody ucznia pełnoletniego na udział w badaniu panelowym oraz przetwarzanie danych osobowych
- Aneks 4a – badanie jakościowe: scenariusz FGI do wywiadów z wszystkimi grupami
- Aneks 4b – badanie jakościowe: scenariusz IDI do wywiadu z uczniem
- Aneks 4c – badanie jakościowe: scenariusz IDI do wywiadu z absolwentem
- Aneks 4d – badanie jakościowe: scenariusz IDI do wywiadu z dyrektorem szkoły
- Aneks 4e – badanie jakościowe: scenariusz IDI do wywiadu z nauczycielem
- Aneks 4f – badanie jakościowe: scenariusz IDI do wywiadu z doradcą zawodowym
- Aneks 4g – badanie jakościowe: scenariusz IDI do wywiadu z pracodawcą
- Aneks 4h – badanie jakościowe: scenariusz IDI do wywiadu z przedstawicielem samorządu terytorialnego na poziomie powiatu