



Krótki opis procesu wprowadzania danych w programie:

EpiData

EpiData jest programem do wprowadzania danych i tworzenia dokumentacji danych.

Użyj programu EpiData, gdy masz zebrane materiały w postaci ankiet papierowych a chcesz tworzyć analizy lub raporty z tych danych. Wprowadzając dane EpiData może zrobić podstawowe tablice częstości, ale program jest nastawiony przede wszystkim na wprowadzanie i dokumentowanie danych.

W czasie wprowadzania danych można od razu je przeliczać, sumować, dokonywać kontroli wprowadzania przez ograniczenie możliwych wartości. Wprowadzając dane, wybieramy je z listy (np. 1=Tak, 2=Nie) a odpowiadające im kody liczbowe są zapisywane w pliku. Wybrane wartości można potem wyeksportować do programów statystycznych jako etykiety wartości zmiennych. Bardzo łatwo wprowadza się daty np. wpisanie „2301” od razu będzie rozumiane przez program jako 23/01/2003 jeśli wprowadzone będą w roku 2003 i w polu „dd/mm/rrrr”.

EpiData jest tak samo użytecznym programem dla zwykłych zbiorów danych (jeden kwestionariusz, lub jeden formularz danych laboratoryjnych), jak i dla wielu połączonych formularzy danych. Główna zasada programu jest zapożyczona w prostym DOSowym programie EpiInfo (w wersji 6), którego od dawna używa wielu ludzi na świecie. Program EpiData pracuje z tą samą strukturą danych i zasad ich tworzenia co EpiInfo6, ale w środowisku Windows i ze szczególnym naciskiem na dokumentację danych.

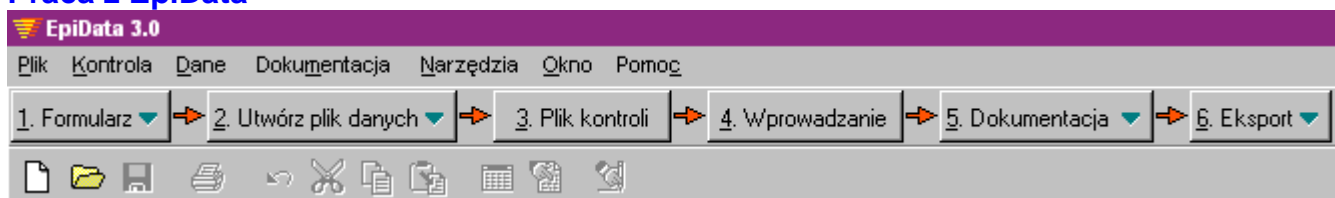
Idea polega na tym, że użytkownik pisze zwykły tekst a program zamienia go na formularz wprowadzania danych. Kiedy formularz jest już gotowy można bardzo łatwo zdefiniować jakie typy danych i jakie zakresy wartości mają być wprowadzane w odpowiednie pola formularza.

Pobieranie i instalowanie programu. EpiData można używać bez żadnych opłat. Pobierz program z www.epidata.dk i postępuj zgodnie z instrukcjami zawartymi w pliku instalacyjnym po uruchomieniu. Program EpiData to jeden plik wykonywalny i pliki pomocy (Mówiąc językiem bardziej technicznym: EpiData składa się z kilku plików, które nie są zależne od żadnych specjalnych bibliotek dynamicznych DLL i nie zmieniają żadnych plików w systemowym folderze komputera. Opcje użytkownika zapisywane są w pliku INI).

Ograniczenia programu

Nie ma żadnych ograniczeń co do ilości zbieranych danych (program testowano z ponad 100 tys. rekordów). Przeszukiwanie ponad 80 tys. rekordów danych indeksowanych zajęło na komputerze z Pentium I 200MHz mniej niż 1 sekundę. Specyfikacja pliku danych musi się zmieścić w 999 wierszach tekstu.

Praca z EpiData



Okno programu EpiData jest standardowym oknem Windows. Posiada jeden pasek menu i dwa paski narzędzi. Pasek „toku pracy” prowadzi od zdefiniowania nowego formularza danych do eksportu danych do dalszej analizy.

1. Budowanie formularza danych

Aby zbudować nowy formularz należy wpisać trzy informacje do każdej zmiennej:

- A.. Nazwa pola (nazwa zmiennej, np.: v1 lub dawka).
- B.. Tekst opisujący zmienną lub pytanie w ankiecie . (np.: „płeć” lub „dzień urodzenia”)
- C.. Definicja rodzaju i długości pola, np.: ## dla 2-cyfrowej liczby.

```

Mój pierwszy formularz wprowadzania danych
id          (automatycznie nadawany numer
identyfikacyjny)          <idnum >
v1          plec          #
v2          Wzrost (meter)  #.##
v3          Masa (kilo)    ###.#
bmi         Body Mass Index ###.##
v4          Data urodzenia <dd/mm/yyyy>
age         Wiek          ###
s1          Kraj          _____
s2          Miasto        <a >
t1         Disiejsza data <Today-dmy>

```

Inne typy pól to Boolean (zerojedynekowe, tak/nie), szyfrowane lub Soundex (tekstowe, „półszyfrowane”)

Ustawienia opcji (menu Plik)

Użytkownik może się zdecydować na jeden z dwóch sposobów nazywania zmiennych:

Opcja „pierwszy wyraz” powoduje, że nazwy pól/zmiennych będą odpowiadać pierwszemu wyrazowi w definicji pola.

Opcja „automatyczne nazywanie” powoduje, że nazwy pól będą się składać z pierwszych 10 liter zdania definiującego pole.

a. v1plec (10 pierwszych liter w zdaniu) - opcja „automatyczne nazywanie”

b. v1 (pierwszy wyraz w zdaniu) – opcja „pierwszy wyraz”

Inne opcje to np.: wybór koloru tła i pola, wielkość odstępu między liniami i inne.

Przykładowo użytkownicy programów Stata lub SPSS powinni używać zasadę nazywania „pierwsze słowo” – nazwy pól będą bowiem nazwami zmiennych. Ponadto użytkownicy Stata powinni używać opcji „Nazwy pól małą literą”.

2. Tworzenie pliku danych.

Po zbudowaniu formularza można podejrzeć wygląd formularza przy wprowadzaniu i utworzyć plik danych.

3. Budowanie kontroli danych podczas wprowadzania

Mocną stroną programu jest możliwość określenia **specjalnych reguł i obliczeń dotyczących wprowadzanych danych**.

- Ograniczenie wprowadzanych wartości zmiennych tylko do określonego zbioru wartości z możliwością wyświetlania tego zbioru podczas wprowadzania.
- Szczególna kolejność wypełniania pól. Np.: wprowadzanie pewnych danych tylko dla mężczyzn (**przeskoki**).
- Obliczenia w czasie wprowadzania. Np.: wiek w czasie wizyty obliczany jako różnica daty wizyty i daty urodzenia. Operacje na skalach i wskaźnikach.
- Wyświetlanie okna z podpowiedzią i inne rozszerzone objaśnienia do obliczeń, procedur warunkowych („if” - „then” - „endif”)

(Obejrzyj przykłady plików kontroli zawarte w instalacji programu, lub pobierz inne przykłady ze strony:

www.epidata.dk/examples.php).

Kiedy uruchomisz edytor pliku kontroli pokaże się nowe okno:

Na górze jest nazwa zmiennej (GIVP02), potem etykieta („typ kwestionariusza givp02:”, poniżej typ pola (Liczba).

Następnie mamy blok definicji zasad kontroli:

Range, Legal definiuje zakres możliwych wartości.

Jumps określa przeskok po wprowadzeniu

(tu mamy zapis mówiący, że po wprowadzeniu 1 program ma przeskoczyć do pola s01vp01a)

Must enter: Jeśli ustawiono wartość na „Yes”, to oznacza, że w tym polu wartość musi być wprowadzona (gdy jest „No” zostawienie pustego pola jest akceptowane)

Repeat: Ustawienie tego parametru na „Yes” powoduje powtarzanie w tym polu wartości z poprzedniego rekordu. np.: gdy wprowadzamy dane uczniów z klasy „Ia” to wartość ta będzie powtarzana przy wszystkich uczniach, aż zaczną się ankiety uczniów z innych klas. Wartość ta może być ciągle zmieniana w trakcie wprowadzania.

Value label: Dla danych określających pewne kategorie możemy zdefiniować odpowiadające im etykiety. Przykładowo 1 może oznaczać „mężczyzna”, a 2 „kobieta”. Wartości te mogą być edytowane przez wybranie „+”. Pole - lista przy tej pozycji pozwala na przyporządkowanie istniejącego bloku etykiet.

Edytuj: Wszystkie inne funkcje można wpisać „ręcznie” po kliknięciu tego przycisku. Zobacz wszystkie dostępne polecenia kontroli w pliku pomocy.

Zapisz: Zapisuje bieżące ustawienia kontroli.

W powyższym przykładzie zarówno warunki **Range,Legal** jak i **Value labels** są określone. zazwyczaj nie ma takiej potrzeby, wystarczy tylko jeden z nich.

GIVP02	
typ kwestionariusza givp02:	
Liczba	
Range, Legal	1-4,9
Jumps	1>s01vp01a
Must enter	Yes
Repeat	No
Value label	etyk_givp02

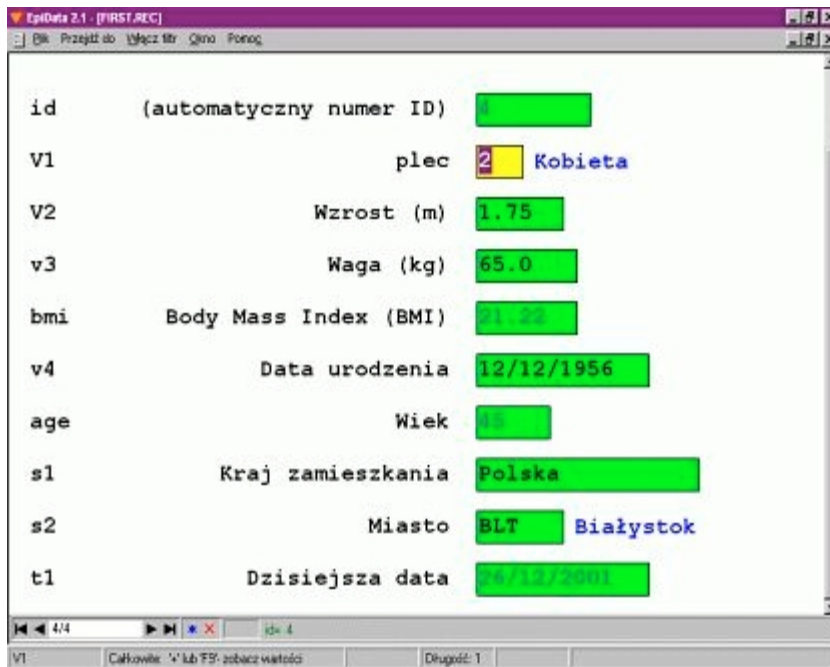
4. Wprowadzanie danych

Otwórz plik i edytuj, wprowadzaj lub przeszukuj dane.

Niebieski tekst w prawej strony pola dodany jest przez program EpiData po wprowadzeniu danych do tego pola i zależy od komend w pliku kontroli. Body Mass Index (BMI) i wiek są obliczone automatycznie.

Używane pliki:

- A. Plik formularza.
Np.: **first.qes**
- B. Plik danych (z rekordami)
. Np.: **first.rec**
- C. Plik z zasadami kontroli danych. Np.:
first.chk
- D. Pliki dodatkowe, np.: **first.not** z notatkami wpisanymi w trakcie wprowadzania danych lub **first.log** z dokumentacją pliku danych.



5. Dokumentowanie danych

Po stworzeniu pliku danych można wydrukować jego strukturę:

Plik danych:	E:\Samples\first.rec					
Opis pliku:	Mój pierwszy plik danych					
Rozmiar pliku:	1129 bajtów					
Zmodyfikowany:	27. lis 2003 12:39					
Liczba pól:	10					
Liczba rekordów:	3					
Parametry kontroli:	Tak (Zmodyfikowany 27. lis 2003 12:25)					
Pól w pliku danych:						
L.p.	Nazwa	Etykieta	Typ	Rozm.	Kontrola	Et. wartości
1	ID	(automatyczny numer ID)	Autonumer	6		
2	V1	plec	Liczba	1		sex 1: Męczyzna 2: Kobieta 9: Nie znana
3	V2	Wzrost (m)	Liczba	4:2	Legal: 0.0-2.30,9	
4	V3	Waga (kg)	Liczba	5:1	Dodatkowe informacje	
5	BMI	Body Mass Index (BMI)	Liczba	5:2	NoEnter	
6	V4	Data urodzenia	Data (dmm)	10	Dodatkowe informacje	
(resztę pól pominięto)						

A po wprowadzeniu danych można wydrukować wartości wszystkich lub niektórych rekordów (można również obejrzeć dane w widoku arkusza z, lub bez etykiet wartości zmiennych):

Rekord 1						
id	1	v1	Męczyzna	v2	1.92	
v4	12/12/1945	s1	Polska	s2	Ciechanów	
t1	28/10/2000					

Spis kodów (codebook) może zawierać również tabelę częstości poszczególnych zmiennych. (przykład nie dotyczy pliku first.rec)

V2	-----	Wzrost (m)
	typ: Liczba(2 po przec.)	

```

kontrola: V2
          RANGE 0.0 2.30
          LEGAL
            9
          END
        END
braki danych: 0/3
zakres: [1.65 ; 1.85]
wartości unikatowe: 3

          średnia: 1.7500
odchylenie std.: 0.0816
V3 ----- Waga (kg)
          typ: Liczba(1 po przec.)
kontrola: V3
          AFTER ENTRY
            BMI=v3/(v2^2)
          END
        END
braki danych: 0/3
zakres: [67.0 ; 100.0]
wartości unikatowe: 3

          średnia: 80.8333
odchylenie std.: 13.9901

```

6. Eksport danych do analizy.

Procedura „kopia zapasowa” tworzy kopie wszystkich plików skojarzonych z danym plikiem danych we wskazanym przez użytkownika katalogu/folderze. Jeśli dane są eksportowane do Stata, to należy pamiętać o małych literach w nazwach zmiennych.

Jednym z nowych narzędzi EpiData jest możliwość stworzenia archiwum ZIP również z opcją szyfrowania z hasłem. Szyfrowanie wykorzystuje algorytm AES/Rijndael akceptowany przez wiele autorytetów w dziedzinie danych komputerowych.

Przy używaniu szyfrowania należy pamiętać, że **nie** ma metody na odzyskanie danych, gdy zgubimy hasło

Inne narzędzia

EpiData zawiera takie funkcje jak porównywanie dwóch plików danych i wylistowanie różnic między nimi. Poprawianie struktury pliku danych po utworzeniu. Kodowanie hierarchiczne, wprowadzanie danych relacyjnych (tworzenie relacyjnej bazy danych), sprawdzanie logicznej spójności danych. Planowane w przyszłości dodatkowe funkcje to np.: możliwość rozszerzenia przez użytkownika systemu menu a także możliwości języka programowania kontroli wprowadzanych danych.

Analizy

EpiData zawiera funkcje opisu całej populacji, jak i podgrup (codebook).

Pracujemy nad stworzeniem modułu do podstawowej analizy danych. Proszę śledzić informacje na www.epidata.dk. Można również użyć eksportu danych do innych programów poprzez proste pliki tekstowe (dane i nazwy zmiennych jako tekst ASCII rozdzielony przecinkami, dBaseIII, lub Excel), a także poprzez całkowity eksport danych, nazw zmiennych, etykiet do programów: Stata, SAS i SPSS.

Gdzie szukać pomocy?

Pliki pomocy i inne dodatkowe dokumenty pomocne w używaniu EpiData można darmowo pobrać ze strony: <http://www.epidata.dk>.

O programie EpiData:

EpiData jest programem opartym na systemie Windows 95/98/NT/2000 (32-bitowym) i służy do wprowadzania danych.

Rozwój programu EpiData został zapoczątkowany przez dr n. med. Jensa M. Lauritsena z Danii

Program jest rozpowszechniany jako darmowy przez Stowarzyszenie EpiData: (The EpiData Association), z Odense w Danii

Program zaprojektowali: Jens M.Lauritsen, Michael Bruus.

Proponowany sposób cytowania:

Lauritsen JM & Bruus M. EpiData (version 3). A comprehensive tool for validated entry and documentation of data. The EpiData Association, Odense, Denmark, 2003.

First version of EpiData released as Lauritsen JM, Bruus M., Myatt MA, EpiData, version 1.0-1.5. A tool for validated entry and documentation of data. County of Funen Denmark and Brixton Health UK. 2001.

Fundatorzy:

Aktualna lista podmiotów wspierających finansowo EpiData jest dostępna pod adresem: www.epidata.dk/funding.htm. Inne informacje na ten temat i podziękowania: [Http://www.epidata.dk/credit.htm](http://www.epidata.dk/credit.htm).

O przekładach programu na inne języki można się dowiedzieć na stronie: [Http://www.epidata.dk](http://www.epidata.dk)

Disclaimer

The EpiData software program was developed and tested to ensure fail-safe entering and documentation of data. We made every possible effort in producing a fail-safe program, but cannot in any circumstance be held responsible for errors, loss of data, work time or other losses incurred by or in relation to the program.

Oświadczenie:

Żeby zapewnić bezpieczne wprowadzanie i dokumentowanie danych, program EpiData został gruntownie przetestowany. Dołożyliśmy wszelkich możliwych starań do tego, by program był wolny od wszelkich błędów, ale nie ponosimy odpowiedzialności za jakiegokolwiek stratę danych, czasu powstałych w wyniku używania programu lub przy jego pomocy.