



Een korte samenvatting over datainvoer met

EpiData

© Jens M. Lauritsen, Michael Bruus & Mark Myatt .

Funen County, Denemarken en Brixton Health, Wales Verenigd Koninkrijk. 2001.

Voor overige informatie and downloaden van de laatste versie: Zie <http://www.epidata.dk>

Wat is EpiData ?

EpiData is een programma voor DataInvoer en dokumentatie van gegevens .

Gebruik EpiData wanneer u gegevens heeft verzameld op papier en daarna een statistische analyse of weergave in tabelvorm wilt verrichten. Er kunnen frequentie tabellen en overzichten van gegevens mee gemaakt, maar EpiData is vooral toegesneden op datainvoer en dokumentatie van gegevens.

Tijdens datainvoer kunnen berekening worden gedaan of restricties aan de in te voeren data worden gesteld. U kunt een waarde kiezen uit een lijst en de corresponderende numerieke code (bijv: 1 = Nee en 2= Ja) opslaan, de tekstlijsten worden geëxporteerd als "waarde labels" voor statistische programma's. Datums invoeren is gemakkelijk: bijv. 2301 wordt opgeslagen als 23/01/2001 wanneer het wordt ingevoerd in het jaar 2001 in een "dd/mm/yyyy" veld.

EpiData is bruikbaar voor eenvoudige gegevensverzamelingen waarbij sprake is van één bron van data (bijv: één vragenlijst of één laboratorium registratie formulier) als wel voor gegevensverzamelingen met veel en/of vertakkende dataformulieren. Het principe is geworteld in de eenvoud van het DOS programma EpiInfo version 6, dat vele gebruikers heeft over de hele wereld. EpiData implementeert de EpiInfo version 6 bestandsstructuur en principes in een windows omgeving met de nadruk op dokumentatie.

Het uitgangspunt is dat u eenvoudig tekstregels schrijft en het programma zet deze dan om in een datainvoer formulier. Als het datainvoer formulier gereed is ,is het vervolgens eenvoudig om te bepalen welke waarden mogen ingevoerd in de verschillende invulvelden.

Aan het eind van dit dokument vind u een stroomdiagram van de werkwijze.

EpiData verandert niets aan de instellingen van uw computer of besturingssysteem.

Het is een essentieel principe van EpiData om niet te interfereren met de instellingen van uw PC. EpiData bestaat uit één programmabestand en enkele helpbestanden.

(Technisch gesproken: EpiData komt als een aantal bestanden en is niet afhankelijk van, of installeert niet ,of vervangt enig DLLbestand in uw systeem directory. Opties worden opgeslagen in de registry).

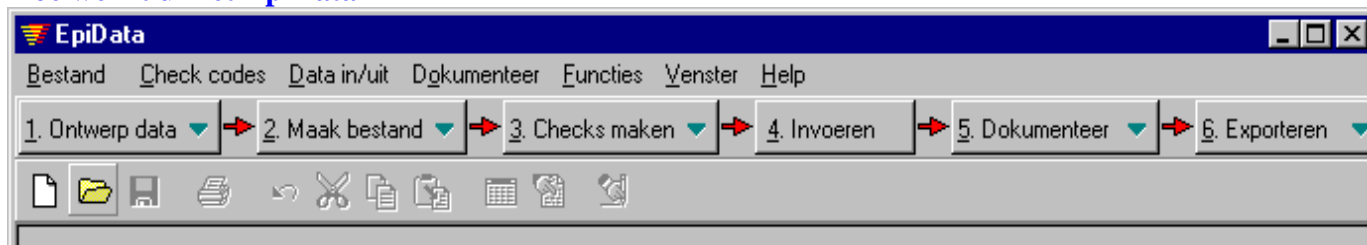
Een standaard "setup.exe" bestand helpt u bij de installatie op uw computer.

Beperkingen

Er zijn geen beperkingen aan het aantal records. (getest met >100.000). Zoeken met een index in 80.000 records duurde < 1 sec op een Pentium I 200Mhz machine).

Alle velden (variabelen) moeten passen binnen 999 regels tekst in het ontwerp.

Hoe werkt u met EpiData



Het EpiData scherm heeft een standaard windows opmaak met één menuregel en twee balken. De "Werkwijze balk" begeleidt u van "1. Ontwerp data" tot "6. Exporteren" voor analyse.

1. Ontwerp Data

U kunt uw dataformulier ontwerpen door drie typen informatie te geven voor iedere variable:

- A.. Naam van het invulveld (variable,bijv: V1 of blootstelling).
- B.. Tekst beschrijving van de variabele. (bijv.: geslacht of "geboortedag")
- C.. Een definitie van het type invulveld, bijv.: ## voor numerieke variabelen met twee cijfers.

```
Mijn eerste Gegevens Formulier
id      (automatisch id nummer) <idnum >
V1      geslacht                #
V2      Lengte (meter)         #.#
V3      Gewicht (kilo)         ###.#
bmi     Quetelett Index        ##.##
V4      Geboortedag            <dd/mm/yyyy>
lft     Leeftijd vandaag       ##.#
S1      In welk landwoonachtig <a >
S2      Woonplaats (Huidig adres) <a >
T1      Invuldatum:           <Today-dmy>
```

Andere veldtypen zijn bijv: boolean (logisch veld, Ja/Nee) of Soundex of tekstveld.

Variabele namen kunnen op twee manieren gemaakt:

- a. V1geslac (de 8 eerste tekens in de zin)
- b. V1 (het eerste woord van de zin).

2. Maak bestand.

Na het schrijven van uw ontwerp kunt u een concept datainvoer formulier inzien of het formulier rechtstreeks maken.

3. Checks maken.

Een sterke kant van EpiData is de mogelijkheid om tijdens het invoeren van data **berekeningen te laten doen** en **invoer restricties te bepalen**:

- Het beperken van gegevens invoer tot zekere waarden en het geven van een beschrijving in tekst van de numerieke codes die worden opgeslagen.
- Bepalen van de volgorde van gegevens invoeren. Bijv: het invullen van gegevens alleen voor manlijke proefpersonen, (**jumps**)
- Toepassen van berekeningen tijdens gegevens invoer. Bijv: leeftijd berekenen gebaseerd op de dag van bezoek en de geboortedatum. Samenvoegen van waarden ,en het indexeren.
- Help meldingen en andere uitgebreidere definities of berekeningen, **if .. then ...endif** structuren voor voorwaardelijke invoer.

(Voor een voorbeeld: haal *eerste.chk* op van: [Http://www.epidata.dk/downloads/examples/](http://www.epidata.dk/downloads/examples/)).

4. Invoeren

Open het bestand en vul waarden in, voeg toe of verander, en zoek gegevens. De kleuren voor de invulvelden en de achtergrond kunnen worden ingesteld. Hier ziet u een lichtgrijze achtergrond met groene velden, het grijze veld is het actieve veld tijdens het invullen.

De verklarende tekst in rood en in blauw rechts van het invulveld maakt EpiData nadat data gebaseerd op labels in het check bestand, zijn gekozen. De Quetelet index en leeftijd worden automatisch berekend.

Opgeslagen bestanden

A. Eerste.qes

Uw dataformulier ontwerp.

B. Eerste.rec

Het huidige databestand bevattende de gegevens.

C. Eerste.chk

Een bestand met daarin de check commando's.

D. Toegevoegde bestanden,

Eerste.not met notities, gemaakt bij de datainvoer of
Eerste.log met documentatie.

Mijn eerste Gegevens Formulier.

ID	Automatisch id nummer	2
V1	Geslacht	1 man
V2	Lengte (meter)	1.80
V3	Gewicht (kilo)	83.0
BMI	Quetelet Index	25.62
V4	Geboortedatum	05/10/1931
LFT	Leeftijd nu	69.5
S1	In welk land woonachtig	PT Portugal
S2	Huidige woonplaats	AMSTERDAM
T1	Invuldatum	17/03/2001

5. Dokumenteer

Nadat u het databestand heeft gemaakt kunt u de bestandsstructuur documenteren.

Een voorbeeld is: (gedeelte van *Eerste.rec*) is:

```

DATABESTAND:      D:\EPIDATA\VERSIE15\Samples\Eerste.rec
Bestandslabel:    Eerste poging
Bestandgrootte:   1048 bytes
Laatste revisie:  17. mrt 2001 16:08
Aantal velden:    10
Aantal records:   2
Checks toegepast: Ja (Laatste revisie 17. mrt 2001 15:29)

Velden in data-bestand:

```

Nr.	Naam	Variabele label	Veldtype	Veldlengte	Checks	Waarde labels
1	ID	Automatisch id nummer	ID-number	6		
2	V1	Geslacht	Integer	1	Must enter	1: man 2: vrouw 3: onbekend
3	V2	Lengte (meter)	Fixed number	4:2	Must enter Legal: 0.5-2.00	
4	V3	Gewicht (kilo)	Fixed number	5:1	Must enter Legal: 1-100	
5	BMI	Quetelet Index	Fixed number	5:2	Meer: Zie Checkbesta	
6	V4	Geboortedatum	Date (dmy)	10	Must enter Legal: 01/01/1910-01/01/2099	
7	LFT	Leeftijd nu	Fixed number	4:1	Must enter Meer: Zie Checkbesta	
8	S1	In welk land woonachtig	Text	5	Must enter	NL: Nederland DE: Duitsland SF: Finland

En na het invullen van gegevens kunt u met *dataoverzicht* waarden van een aantal of alle records in tabelvorm zien.

Lijst met records in D:\EPIDATA\VERSIE15\Samples\Eerste.rec
Lijst gemaakt: 18. mrt. 2001 16:30

Observatie 1 (record #1)

ID	1	V1	vrouw	V2	1.67
V3	61.2	BMI	21.94	V4	20/08/1955
LFT	45.6	S1	Nederland	S2	PURMEREND
T1	17/03/2001				

Observatie 2 (record #2)

ID	2	V1	man	V2	1.80
V3	83.0	BMI	25.62	V4	05/10/1931
LFT	69.4	S1	Portugal	S2	AMSTERDAM
T1	17/03/2001				

Een codeboek afbeelding kan ruwe frequentie tabellen bevatten. (voorbeeld gebaseerd op Eerste.rec en slechts weergave van 2 variabelen)

```

CODEBOEK
=====
Rapport gemaakt          18. mrt 2001 16:56
Databestand:            D:\EPIDATA\VERSIE15\Samples\Eerste.rec
Bestands_label:         Eerste poging
Bestands_datum:         17. mrt 2001 16:08
Checks toegepast:       Ja (Laatste revisie 17. mrt 2001 15:29)
Aantal velden:          10
Records totaal:         2
Verwijderde records:    0
Gebruikt in codeboek:   2 records

ID ----- Automatisch id nummer
          Type: ID-number
          Niet aanwezig: 0/2
          unieke waarden: 2

V1 ----- Geslacht
          Type: Integer
          Waardelabels: Onbenoemde waardelabel set
          Niet aanwezig: 0/2
          Omvang: [1,2]
          unieke waarden: 2
          Tabelweergave: Frequentie Procent Waarde Label
                          1 50.0 1 man
                          1 50.0 2 vrouw

```

6. Exporteren (voor analyse en om de gegevens veilig op te slaan.)

De backup functie kan alle bestanden die samenhangen met een bepaald gegevensbestand ,kopieren naar een door de gebruiker tevoren gemaakte directory /folder die dient voor de backup.

Funcities

Epidata bevat nog ander mogelijkheden zoals:

- 1: Het vergelijken van twee bestanden en het opsommen van verschillen tussen records.
- 2: Het kunnen bijwerken van de structuur van het gegevens bestand.
- 3: Hierarchische codering in de checkcodes.
- 4: Tenslotte wanneer het programma geheel gereed is, ook relationele datainvoer, beoordeling van logische consistentie en functies waarmee door de gebruikers gemaakte uitbreidingen aan het menusysteem en aan de checkcode-taal kunnen toegevoegd.

Analyse van de gegevens

De bestanden die gemaakt zijn in EpiData kunnen zonder meer geanalyseerd in EpiInfo 6., in de module Analysis.exe Voor andere programma's is het mogelijk data te exporteren als bestand.

-- Data en variabele namen: comma gescheiden ascii bestand, (.txt)

-- dBaseIII bestand (.dbf)

-- Excel bestand (.xls)

of een complete export (data en variabele namen, plus labels voor Stata versie 4-6; en later ook SPSS).

Mogelijkheden voor ondersteuning

Het helpbestand en andere aanvullende documenten zijn vrij beschikbaar van <http://www.epidata.dk>.

Omtrent EpiData

EpiData is een Windows 95/98/NT/2000 gebaseerd programma (32 bit) voor DataInvoer.

Programma ontwerp door Jens M. Lauritsen, County of Funen, Denemarken., en Michael Bruus, Denemarken , en Mark Myatt, Brixton Health, Wales en The Institute of Ophthalmology, London, UK.

Programmering door Michael Bruus, Denemarken.

Fondsen en dankzegging.

We willen dank zeggen voor de financiële bijdragen aan de ontwikkeling van **EpiData** door: Provincie Funen, Denemarken. Brixton Health, Wales UK. Deense Data Archiven/ERAS, Denemarken, Universiteit van Zuid Denemarken, Faculteit van Gezondheid. Valid International. London UK. Londen School voor Hygiene & Tropische Geneeskunde, UK. International Centre for Eye Health, Londen UK.

EpiData werd zodanig ontworpen dat de qes-chk-rec bestand structuur zoals in Epi Info v.6 gehandhaafd bleef.

Epi Info 2000 volgt een ander weg. U kunt meer te weten komen over Epi Info 2000 op de

Epi Info website, zie <http://www.cdc.gov/epiinfo>. Voor een complete lijst van dankzeggingen zie het help

bestand. In het bijzonder zijn wij erkentelijk iedereen die hielp de EpiInfo v.6 principes te ontwikkelen.

Internationale versies werden gemaakt door anderen. Zie ook de eerste pagina van dit document. Voor

bijdragen of donaties voor verdere ontwikkeling en voor de documentatie kunt u een

e-mail sturen aan: info@epidata.dk

Voorgestelde aanhaling:

Lauritsen JM, Bruus M, Myatt M. EpiData, A tool for validated dataentry and documentation of data. County of Funen Denmark & Brixton Health UK. 2000. (version x.y). Voor vertaalde versies svp toevoegen

:(Nederlandse taal versie vertaald door Louk Meertens)

Aansprakelijkheids beperking

Het EpiData software programma werd ontwikkeld en getest om foutloos datainvoer en documentatie te kunnen verrichten. Wij hebben alle mogelijke inspanning gedaan om een veilig en foutloos programma te maken, maar kunnen in geen enkel geval verantwoordelijk worden gehouden voor het verlies van data, werktijd of andere schade, opgelopen door of in relatie met, dit programma.