



Kratek pregled vnosa podatkov s pomočjo

EpiData

EpiData je program z vnos in za dokumentiranje podatkov.

Program EpiData uporabite kadar ste podatke že zbrali in jih želite statistično analizirati ali jih tabelirati. Možno je narediti osnovne tabele pogostnosti in sezname toda je EpiData usmerjen na vnos podatkov in na dokumentiranje le-teh.

Možno je vnaprej nastaviti, da se v teku vnosa podatkov izračunajo sumarne vrednosti ali pa definirati omejitve vnosa vrednosti. Lahko si iz seznama izberete odgovarjajočo numerično kodo (1 = Ne 2 = Da), možno je eksportiranje seznamov v obliki teksta kot "value labels-oznake vrednosti" kar se nato uporablja v statističnih programih. Vnos datumov je olajšan, npr. 2301 bo oblikovano kot 23/01/2003, če v polje "dd/mm/llll vnesemo le vrednost letnice 2003.

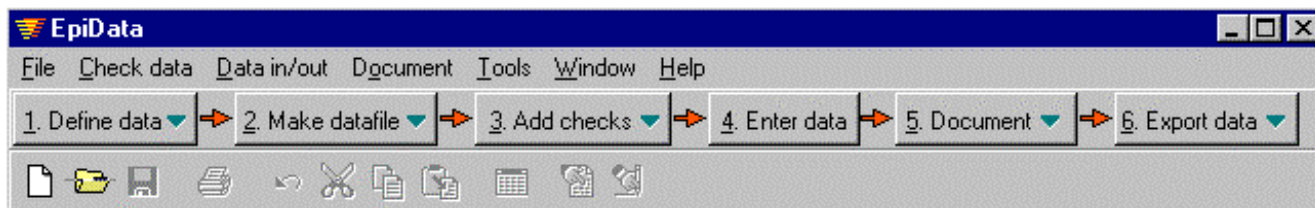
EpiData je zelo uporaben za enostavne podatkovne zbirke kadar imamo eden vir podatkov (npr. eden vprašalnik ali eno vrsto laboratorijskega izvida). Enako tako je uporaben za vnos podatkov iz številnih formularjev, ki so razvejani v tem dokumentu je opisana le enostavna situacija. Principi tega programa izvirajo iz enostavnosti dos programa Epi Info v6. Ideja v tem, da uporabniku omogoči enostavno pisanje vrstic teksta, ki jih bo program konvertiral v obrazec-vprašalnik za vnos podatkov v katerem je možno oblikovati in dodati kontrolo vnosa, pogojne skoke v polja z vnos ali izračune vrednosti v poljih.

Downloading in inštaliranje EpiData je zastoj program. Lahko ga posnamete -download z spletnega naslova www.epidata.dk in ga inštalirate spremljajoč inštrukcije. EpiData ne bo motil dela vašega računalnika. EpiData se sestoji iz ene programske datoteke in iz help datotek. (Povedano s tehničnimi termini: EpiData se obnaša kot nova datoteka, ki ni odvisna, se ne inštalira niti ne zamenja katerokoli DLL datoteko v vaši sistemski direktoriji. Opcije se shranijo v posebni ini datoteki).

Omejitve

Ni omejitev števila observacij. (testiran je z >100.000). Search-iskanje z indeksom v 80.000 zapisov < 1 sec na Pentium I 200Mhz). Specifikacija strukture podatkovne datoteke mora biti znotraj limita 999 vrstic teksta.

Kako dealti z EpiData



Zaslona programa EpiData se prikaže v "standardni" windows obliki, ki vsebuje 1 vrstico menijev in dve orodni vrstici.

Orodna vrstica "Work Process toolbar-delovni proces" nas pelje od "1. Definiranje podatkov" do "6. Eksport podatkov" namenjenem analizi le-teh.

1. Definiranje podatkov

Podatke definiramo z vpisovanjem 3 vrst informacij za vsako spremenljivko:

- Naziv polja z vnos(spremenljivka, tj. v1 ali ekspozicija).
- Tekst, ki opisuje spremenljivko. (npr. spol ali "datum rojstva")
- Definicija polja, npr. ## za dvomestno številčno polje.

Ostala polja so boolean (da-ne), encryptirana ali soundex polja.

```
My first DataEntry Form
id          (automatic id number)  <idnum >
V1          sex                    #
V2          Height (meter)         ###
V3          Weight (kilo)          ###.#
bmi         Body Mass Index        ###.#
v4          Date of birth          <dd/mm/yyyy>
age         Age today              ###
s1          Country of Residence
s2          City (Current address)  <a >
t1         Today's Date            <Today-dmy>
```

Nastavitev opcij(meni datoteka) Pomembno je, da

uporabnik določi edan ali dva principa na osnovi katerih bo imenoval spremenljivke. Pri opciji "first word-prva beseda je naziv" program določi naziv polju s prvo besedo, ki jo vpišemo v vrstici. Pri opcijo "automatic-avtomatično" program za naziv polja določi prvih 10 znakov v vrstici. Primeri:

- | | |
|--------------------------------------|-----------------------------------|
| a. v1spol (10 prvih znakov v stavku) | "automatic" naziv polja je v1spol |
| b. v1 (prva beseda v stavku). | "prva beseda" naziv polja je v1 |

Ostale opcije: npr. **barva ozadja in barva polj, višina vrstice itd.** Uporabniki programov kot so npr. Stata ali SPSS naj uporabljajo princip "prva beseda"; nazivi polj bodo nazivi spremenljivk. Uporabniki programa Stata morajo za nazive polj uporabiti male črke.

2. Naredi datoteko s podatki.

Po vpisovanju definicij lahko naredimo predogled našega obrazca ali pa oblikujemo datoteko s podatki.

3. Dodaj/Revidiraj kontrole-Checks – ob Vnosu podatkov

Mogočen del EpiData is je možnost, da **specificiramo pravila in izračunavanja tekom vnosa podatkov.**

- Lahko omejimo vnos na samo določene vrednosti in podamo opis numeričnih šifer, ki se vnašajo.
- Specificiramo zaporedje vnosa podatkov. Npr. vnašamo odgovore na vprašanja, ki se nanašajo le na moške, (**skoki-jumps**)
- Oblikujemo ukaze za izračunavanje tekom vnosa podatkov. Npr. starost ob obisku, ki temeljina datumi obiska in na datumu rojstva. Toda najmočnejši izračuni so tekom faze analize podatkov.
- Help sporočila in druge razširjene definicije izračunov, ukazi if .. then ...endif.
(Glej primere, ki so pristotni ob inšatiranem programu EpiData, ali prevzami še primerov na spletu www.epidata.dk/examples.php).

Kadar poženemo opcijo "dodaj/revidiraj" se nam odpre novi zaslon: Na vrhu je naziv spremenljivke (v8) kot je na sliki, pod nazivom je oznaka (Rigidfix) in vrsta spremenljivke (number). Sledijo bloki za definiranje:

Range-raspon, Legal-dovoljeno definira katere podatke je možno vnašati.

Jumps-skoki specificira kam kazalec gre po vnosu podatka (tu: pri vrednosti 1 jump –skoči v polje v10)

Must enter-Obvezen vnos: Če opsijo damo na Da se mora vnesti vrednost. (drugače zapiusti polje prazno)

Repeat-ponovi: Ponovi vrednost, ki je bila vnešena v prejšnjem zapisu, npr. če so podatki iz skupine, ki ima enako vrednost. Vrednost je možno spremeniti.

Value label-oznaka vrednosti: Za kategorične podatke se s tem definira kaj vrednost pomeni. Npr. 1=moški 2=ženska. Vrednosti urejamo s pomočjo "+". Padajoči seznam nam omogoča izbiro definirane oznake vrednosti.

Edit-uredi: Možno je definirati še marsikaj. Glej v help datoteki zbirko ukazov.

Save-shrani: Sshrani podane definicije.

V primeru so definirani **Range-razpon, Legal-dovoljene vrednosti** in **Value labels –oznake vrednosti**. Pri navadnem vnosu podatkov to ni tako. Uporabimo le eno iz the opcij.

4. Vnos podatkov

Odre se datoteka in se jih vnaša lai se jih preiskuje.

Pojasnitveni tekst v modri barvi zraven vrednosti v poljih je doadal program EpiData po vnosu podatkov na osnovi oznak v check datoteki. Primer je izračunana vrednost v poljih Body mass index in age-starost.

Datoteke, ki se shranijo:

- Obrazec za podatke-vprašalnik z definicijami. Npr. **first.qes**
- Aktualna datoteka s podatki, ki vsebuje podatke. Npr. **first.rec**.
- Datoteka z definiranimi kontrolami. Npr. **first.chk**
- Dodatne datoteke, npr. **first.not** z opombami, ki so dodane tekom vnosa podatkov **first.log** z dokumentacijo.

ID	(automatic id number)	1
U1	sex	2 Female
U2	Height (meter)	1.75
U3	Weight (kilo)	67.0
BMI	Body Mass Index	21.88
U4	Date of birth	12/12/1956
AGE	Age today	44
S1	Country of Residence	Denmark
S2	City (Current address)	00 Odense
T1	Todays Date	27/01/2001

5. Documentiranje podatkov

Po oblikovanju datoteke s podatki nam je omogočeno dokumentiranje strukture datotek. Primer (del datoteke *first.rec*):

```
DATOTEKA: C:\data\first.rec
Oznaka datoteke: Moja prva testna datoteka je primer

Velikos:          612 bytes
Zadnja revizija:   28. Jan 2001 12:14
Število polj:     7
Število zapisov:  0
kontrolni ukazi - Checks:   Yes (Last revision 28. Jan 2001 12:02)
```

Polja v podatovni datoteki:

No.	Naziv oznaka	spremenljivke	Fieldtype	Width	Checks	Value labels
1	id	ID-number	6			
2	v1	spol	Integer	1		spol 1: Moški 2: Ženska 9: Nepoznano
3	v2	Višina(meter)	Fixed number	4:2	Legal: 0.0-2.30,9	
6	v4	Datum rojstva	Date (dmy)	10		

(other fields omitted)

Po končanem vnosu podatkov lahko naredimo seznam vrednosti za nekaj ali za vse zapise in lahko pregledamo podatke z ali brez oznak v obliki tabele:

Observation 1

id	1	v1	Male	v2	1.92
v4	12/12/1945	s1	denmark	s2	Copenhagen

"codebook" lahko zajema surove tablice frekvenc. (primer ni iz datoteke *first.rec*)

```
v2 ----- Sex
      type: Integer
value labels: sex
range/legal: 1-2,2
missing: 0/25
range: [1,2]
unique values: 2
tabulation:
      Freq.  Pct.  Value  Label
          11  44.0   1     Male
          14  56.0   2     Female
v3 ----- Temp
      type: Floating point
range/legal: 36.00-40.00
missing: 0/25
range: [36.00,37.50]
unique values: 12
mean: 36,84
std. dev: 0,37
```

6. Ekportiranje zaradi analize in zavarovanje podatkov.

Eksportiramo podatke zaradi analize ali varnostnih kopij-backup v izbrano direktorijo. Kadar podatke eksportiramo za program Stata moramo izbrati opcijo majhne črke.

Kot del menija z orodji obstaja opcija uporabe in stiskanja datotek –zipz opcijo enkripcije le-teh. Encipcija uporablja AES/Rijndael enkripcije ki so se pokazale kot zelo dobre.

Opozarjamo, da **ni možnosti** ponovno najti password, če ste ga pozabili.

[Orodja in ostalo](#)

EpiData vsebuje še druge aspekte kot je primerjanje dveh datotek in izdelava seznama razlik med njimi na nivoju naziva polj. Revidiranje strukture datoteke s podatki brez izgubljanja že vnešenih podakov. Hierarhično šifriranje, vnos v relacijske datoteke, check-kontrola konsistenčnosti logike-logična kontrola. V prihodnosti je v načrtu dodati še funkcije za dodajanje ekstenzij, ki jih poda uporabnik in programski jezik za check datoteke.

Analysis

EpiData zajema funkcije za splošno tabeliranje (codebook). Začeli smo delo na razvoju osnovnega paketa za analize podatkov. Treba je slediti spremembe na spletu www.epidata.dk. Eksportiranje podatkov zaradi analize je možno v različne oblike datotek (nazivi podatov in spremenljivk: comma ločeni v ascii datoteke, dbaseIII ali Excel) ali celoten eksport (nazivi podatkov in spremenljivk, plus oznake: Stata, SAS in SPSS).

Viri za podporo

Help datoteke in različni drugi dodatni dokumenti so dosegljivi na <http://www.epidata.dk>

Diskusijska skupina in podpora soo aktivni že od jeseni l. 2003.

About EpiData

EpiData je na Windows 95/98/NT/2000 bazirani program (32 bit) za vnos podatkov. Razvoj EpiData so inicirali Jens M. Lauritsen, MD. PhD, Denmark. Program izdaja kot freeware : The EpiData Association, Odense Denmark. Program design: Jens M.Lauritsen & Michael Bruus. **(Translated by Marko Kovacevic)**

Predlagano citiranje:

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.

Financiranje in zahvale.

Ažurirani seznam finančne podpore je na [Http://www.epidata.dk/funding.htm](http://www.epidata.dk/funding.htm) ter zahvale na : [Http://www.epidata.dk/credit.htm](http://www.epidata.dk/credit.htm) . Preveden je na več jezikov, glej [Http://www.epidata.dk](http://www.epidata.dk)

Za donacije in za pomoč glej help datoteko ali pošlji e-mail to info@epidata.dk

Opozorilo:

Ob vseh prizadevanjih, da izdelamo varen in zanesljiv program ne moremo pod nobenih okoliščinah prevzeti odgovornosti za napake, izgubo podatkov, delovnega časa ali drugih izgub, ki se lahko pripišejo ali so povezan s programom.