

INFO411 – Exercise data set

Purpose to to learn:

- importance of understanding the data/domain
- working with fixed field data sets
- working with data that contain errors, outliers, noise
- working with dates

Extract from patient transaction records.

Collection of 4,879,152 records from the first quarter of 1999.

All fields are space delimited, terminated by newline, and record length is 60 bytes.

Encrypted fields (flagged as such) are hex strings, however to avoid ECBDIC !ASCII conversion problems these fields be read MUST BE READ as 16 byte character strings.

Field	Variable title	Format
1	Card number (encrypted)	Hex 16
	Current patients medicare card number	
2	Card issue number	Num 1
3	Pin (encrypted)	Hex 16
	Patient Identification Number	
4	Patient gender	Char
5	Patient date of birth	ddmmmCCyy
6	Patient enrolment postcode	Num 4
7	Patient homecode	Num 2
	Country code for the passport in which the person entered Australia	
8	Enrolment type	Char 1
	See attachment S	
9	Payee	Char 1
	Method of payment	
	1= pay claimant	
	2=pay doctor via claimant	
	3=cash payment	
	4>manual cheque payment	
	5=pay claimant, claimant not cardholder	
	6=direct bill	
	8=private Health	
	9=EFT (direct to bank account)	
	E=reject EFT	
	F=reject for Private health	
	R=reject cheque for Patient claim	
	T=reject cash for patient claim	
	S=reject direct Bill	
	X= substantiation of patient contribution	

ATTACHMENT S - ENROLMENT TYPES

CODE	DESCRIPTION
space	Default initial enrolment
A	Australian
B	Academic and Teaching Staff (Higher Education Funding Act 1988)

D		Change of Status (COS), Processing Entry Permits (PEP),
E	Diplomat	Migrant
F		Fulbright Scholars Fund (Australian/American Education
G	Foundation)	Refugee status
L		Employed in Research with Federal Government Funding.
N		Employer Nomination Scheme
O		Employed in Skills Transfer Scheme, Visa 412, 413 or 414.
P		Provisional Resident Visa.
R		RHCA (Reciprocal Health Care Agreement)
S		Overseas Subsidised Student or AIDAB (Australian
T	International	Development Assistance Bureaux)
Z		Temporary Visitor, PRC Nationals Visa Codes 437 and 783
		Refugee Applicant, not a refugee yet.