MODULE 7

DATA INPUT METHODS

WORKED EXAMPLES

7.1 What is the purpose of data validation program?

To detect errors which may have been made by a data entry operator in entering data from forms into a computer's secondary memory. The detected errors are then corrected to ensure that data file has no errors.

7.2 What are the main principles used in designing forms for data entry?

(i) Reduce human efforts in filling forms

(ii) Minimize possibility of errors in entering data from forms into a computer's secondary memory

(iii) Minimize effort in entering data from forms into a computer's secondary memory.

7.3 Design a form to be used by a salesman to report to the office about the sales executed by him at different customer locations

See Table below

A Salesperson Form

Sales Person

Your name :	M	.	R	A	M	A	M	U	R	T	H	Y			

Your code :	M	R	4

Your budget code:	1	2

Sales details

Date:	D	D	M	M	Y	Y

Item Code	Description	Qty. Sold	Price
K248	Toilet soap	256	3.50
J468	Detergent cake	468	2.25
P764	Liquid soap bottles	28	8.45
Enter Totals		752	14.20

Customer details

Customer name:

Customer code:

Customer address:

PIN

Mode of payment (Tick code)

1	CASH	2	CHEQUE	3	BILL

7.4 Is concise code comprehensive? If not, why?

No. In a concise code the aim is to keep the length of the code small whereas in a comprehensive code it is to include as much information as possible about the entity being coded.

7.5 Is meaningful code necessarily comprehensive?

No. Meaningful code aids in recognizing the entity being coded whereas a comprehensive code tries to include as much information as possible about the entity being coded. For example BICYCLE 24 indicates a 24 inch height cycle. It is meaningful. A code such as

BC 24 G R HERO 2684

Describes 24-inch bicycle, which is for gents, red in color, manufactured by Hero with serial number 2684.

7.6 What is the advantage of serial number code? Why is it not normally used?

It is concise, expandable and precise. It is not meaningful or comprehensive and thus not often used.

7.7 Design a group classification code to code (i) motor vehicles, (ii) music cassettes, and (iii) books.

(i) Motor vehicles

Types of vehicle	*Year of manufacture*	*Engine CC*	*Brand*	*Serial no.*
2 alphabets	*4 digits*	*4 digits*	*3 alphabets*	*7 digits*

Mnemonic codes

Vehicle types	*code*
Two wheeler	TW
Three wheeler	RW
Private car	PC
Taxi	TA
Bus	BU

Manufacturer	*Code*
Bajaj	BAJ
Telco	TEL
Leyland	LEL
Maruti	MAR
Ambassador	AMB
Vespa	VSP
Fiat	FAT
Honda	HDA

Sample code: PC 19 88 800 MAR 0056789

(ii) Music cassettes

Types of music	*Nature of music*	*Type*	*Publisher code*	*Serial no.*
2 alphabets	2 alphabets	1 digit	3 digits	4 digits

Mnemonic codes

Music type	*Code*
Classical North	CN
Classical South	CS
Classical West	CW
Film North	FN
Film South	FS
Film West	FW
Pop North	PN
Pop South	PS
Pop West	PW

Nature of music	*Code*
Instrumental	IN
Vocal	VO
Orchestra	OR
Chorus	CH

Type	*Code*
Mono	1
Stereo	2

Publisher code: 3 digits, Serial no.: 4 digits.
Example: FNIN14506784

(iii) Books

ISBN code is a good example.

Area	*Publisher*	*Book no.*	*Check digit*
U.S.	code	8 digits (Total)	
U.K.			
Germany, U.S.S.R. etc.	Publishers publishing a small number of books have a long publisher code and smaller no. of digits for book no.		
Example: 0	| 87692	| 617	| 0

7.8 Give an example of significant code. Are significant codes expandable?

Code for a shirt

SH	40	95	58
↓	↓	↓	↓
Shirt	Collar size (Cm)	Chest size (Cm)	Sleeve size (Cm)

Yes.

7.9 Add a modulus-11 check digit to the codes (i) 48467 (ii) 96432 and (iii) 87646257.

Check digits are respectively

(i) 9 (ii) 8 (iii) 3.

7.10 Modulus–37 check is suitable for alphanumeric codes. Add a modulus-37 character to the codes (i) 4AB9W (ii) XBY483 and (iii) CAZ4642.

(i) 2 (ii) N (iii) N.

7.11 If modulus-11 check digit system is to generate detection of multiple identical digit transcription error (i.e., a code such as 45565 is wrongly entered as 48868), what should be constraints on the weights?

A digit t becomes x for weights w_q, w_r, w_s.

Let w_i be the weights.

$$\sum_{i=1}^{n} w_i d_i = p.N \text{ if no error}$$

$$\sum_{i=1}^{n} w_i d_i = \sum_{i=1}^{n} w_i d_i + (w_q + w_r + w_s)t = p.N$$

$i \neq q,r,s$

The condition for detecting error is

$(w_q + w_r + w_s)(x - t) \not\equiv p.N$

Therefore,

$(w_q + w_r + w_s) \not\equiv p.N$

Sum of any subset of weights should not be equal to 11 or a multiple of 11. Possible only for codes less than 4 digits long, including check digit.

7.12 A see-saw error is one in which one digit of the code is increased by x and another decreased by x. For example, 486732 becoming 456762. When can modulus-N check detect such errors?

Let the k^{th} digit become $(d_k + t)$ and q^{th} digit $(d_q - t)$

Weighted sum = $\sum_{i=1}^{n} w_i d_i + w_k t - w_q t$ with these errors condition is

$\sum_{i=1}^{n} w_i d_i + t(w_k - w_q) \not\equiv p.N$

Satisfied if (i) $w_k \not\equiv w_q$

(ii) N is prime

(iii) $|w_k - w_q| < N$

These are satisfied if $w_k \not\equiv w_q$, $w_k, w_q > 0$ and $w_k, w_q < N$.

Therefore all weights are distinct

7.13 Why is it useful to assign sequence numbers for data records? What are the types of errors detected by sequence numbering?

Can trace missing records using sequence numbers. Records out of sequence can be detected. Excess records (with duplicate sequence numbers) can be detected.

7.14 A set of data records for student examination results has the following format:

Roll No.	*Name*	*Marks (out of 100)*			
		Paper 1	Paper 2	Paper 3	Paper 4

Design for these records a batch control record and a record control field and other appropriate checks for the fields

Batch control record

(i) No. of records in batch
(ii) Sum of marks in papers 1 to 4
(iii) No. of records with marks in papers 1 to 4 >=60
(iv) No. of records with marks in papers 1 to 4 < 40

Record check (i) Sum of marks in papers 1 to 4, (ii) No. of papers in record with marks >= 60

Other checks for the fields

(i) Modulus-11 check for Roll no.
(ii) Each marks field <= 100, >= 0
(iii) Sum field <= 400
(iv) Flag record with marks in any paper >= 80 and another paper <= 30.

7.15 What is the difference between range check and radix check?

Range gives maximum allowable value for a field as determined by the analyst. For example in one paper if maximum marks is 50, range check will use 50 and if it is 100 in another it will use 100 as range. Radix is however an invariant. No. of hours/day are always 24 and is universally known.

7.16 Give some examples of fields where reasonableness check would be applicable.

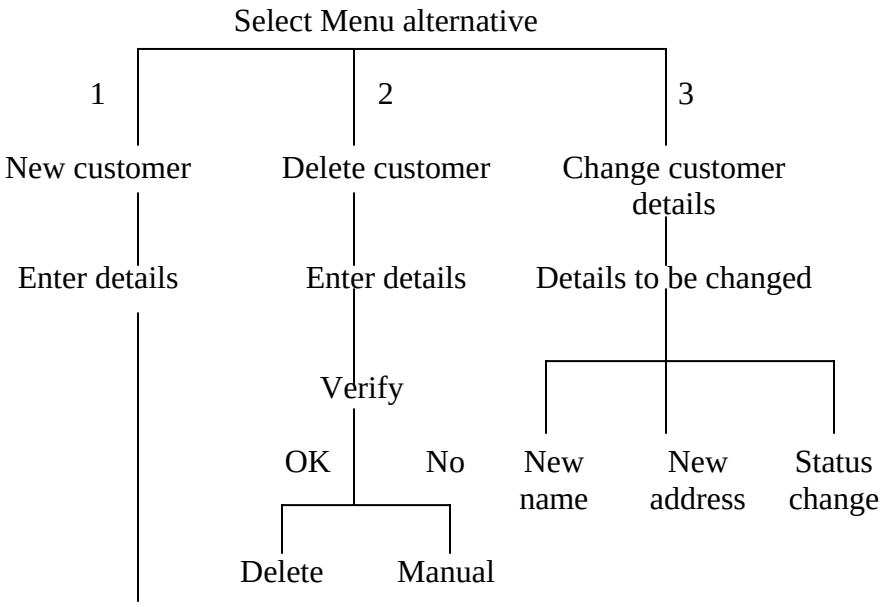
If normal electricity consumption of a consumer is 250 kWh/month, a value of 1500 kWh in a month will be considered unreasonable. Other examples are:

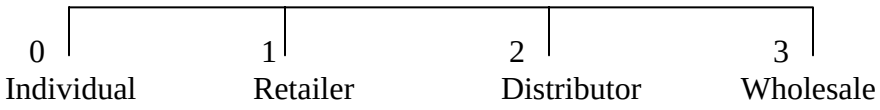
(i) Deductions in a paybill
(ii) Price/unit of some items
(iii) Qty. ordered in comparison to normal averages.

7.17 Give some examples of inter-field relationship checks

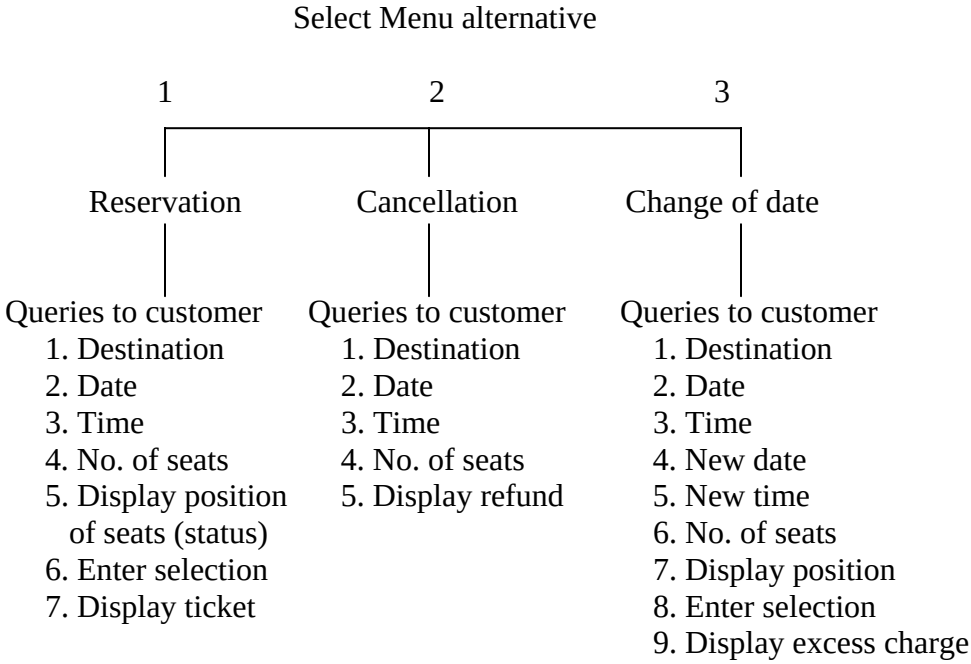
Employee status vs. salary
Age vs. marital status (Age <= 12 cannot normally be married)
Age vs. Education

7.18 Design a dialogue hierarchy for entering data on customers (of a manufacturer).





7.19 Design a dialogue hierarchy and the screens for a system used to reserve seats in long distance buses.



Screen 1

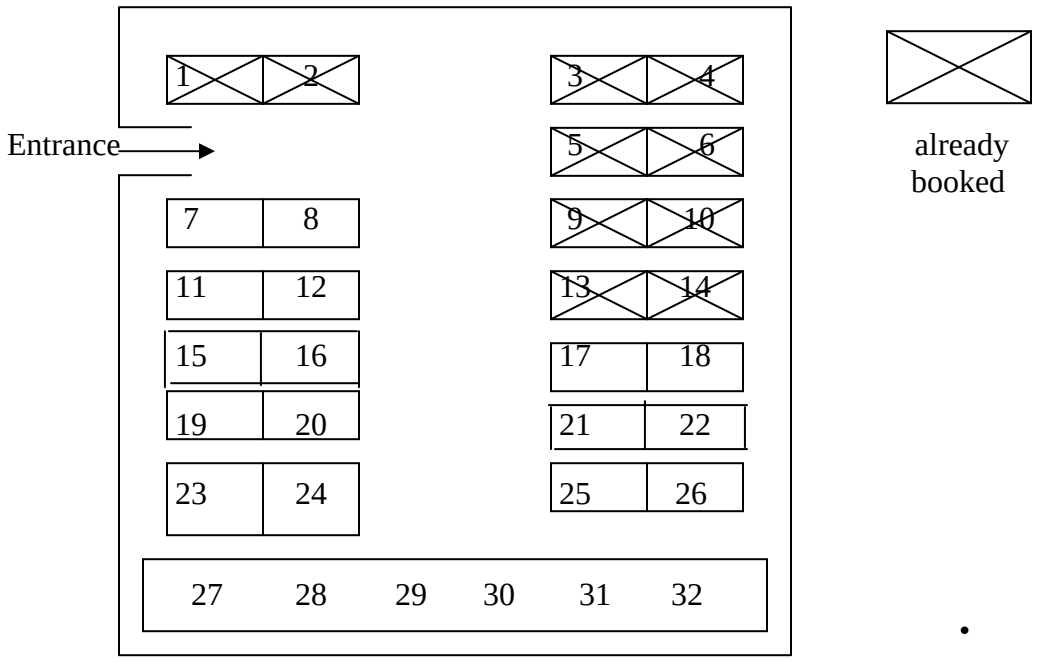
Reservation [1] Cancellation [2] Change of date [3]

Screen 2 reservation commands

- What is your destination ?
→ Mercara
- Date of journey?
→ 260589
- Time of bus
→ 0830
- How many seats?
→ • Adults 2 • child 1
- Screen display Front

- Queries on terminal
  User responses

(Seats in bus)



Which seats do you want ?

→ 19 20 21

- Ticket display

DESTINATION	DATE	TIME	SEAT Nos.	FARE
MERCARA	260589	0830	19 20 21	Rs. 140