

Zair/005/0057/15

nd

ZIMCO INSTITUTE OF MANAGEMENT

INFORMATION TECHNOLOGY

PROCESSING METHODS: DATA VALIDATION

1. BASIC TYPES OF PROCESSING WITH COMPUTERS

There are basically four types of processing methods used in computers.

- a) Batch Input and Batch processing of all data.
 - all transaction records for file updating or file maintenance are batched for input and processing at the same time
 - all file enquiries are input as transaction records and batched for input at the same time.
- b) Batch Input and Batch processing of all transaction data, but with on-line file enquiries.
 - all transaction records for file updating or file maintenance are batched for input and processing at the same time.
 - file enquiries can be made "ad hoc" and individually by means of on-line direct access. The master file must be on a direct access medium (i.e. disc) and kept permanently on-line. It should be organised indexed sequentially, so that it can cater for both batch updating and direct access enquiry.
- c) On-line - Input, batch processing, and on-line enquiry.
 - data is input via remote terminals and stored on a computer file without being processed.
 - the Input data on the temporary transaction file will be periodically used to update the master file in batch mode.
 - file enquiries can be made ad hoc and individually by means of on-line direct access from a terminal. The master must be on a direct access medium and kept permanently on-line.
- d) On-line Input, on-line (real time) updating and on-line enquiry.
 - data is Input from remote terminals
 - the master file is updated when the data is input i.e. at once.
 - file enquiries can be made ad hoc and when required.

This is known as real-time processing, and updating the master file immediately is a central feature of it. The master file must be at a direct access medium and on-line to the CPU at all times.

2. CONTROL OVER DATA COLLECTION AND INPUT-BATCH PROCESSING

As in any other processing methods, the first step is to try and make sure, as much as possible, that the data which will be Input to the computer system is complete and accurate. Batch controls will be established before the documents leave the user department. Each batch will be accompanied by a batch control slip which will include the following data:-

- a) - the batch number
- b) - the computer system identification code
- c) - control totals:

- i. number of records in the batch
- ii. value of transaction in the batch
- iii. hash totals (e.g. total of quantities; total of invoice numbers).
- iv. a combination of more than one of these

- d) - authorisation (date prepared; prepared by; keyed in by; verified by etc).

A hash total is a number which has no significance apart from its use as a data control total.

The batch control data is keyed in with the source document data and throughout the processing of the batch the control information will be checked to ensure that all the correct data and only the correct data is being processed.

3. BATCH PROCESSING EXAMPLE

Chumba Ltd operates a computer based sales ledger. The main stages of processing are:

- a) The sales invoices are prepared manually and one copy of each is retained. At end of day all invoices are clipped together and a batch slip is attached. The sales clerk allocates the next unused batch number in the batch control book. He enters on the control slip, total number of documents, batch number, total value of documents, batch number, total value of invoices.
- b) The batch of invoices is then passed to the data processing department. The control clerk ticks off the batch number received.
- c) The invoice and control details are punched onto cards, or keyed onto magnetic tape, verified, and input into the computer. The first program is a data vet program which performs various edit checks and produces a listing of all valid invoices with document count and total value. It is also produces rejection listings.

- d) The control clerk reconciles the totals on the batch control slip with the totals for valid and rejected data.
- e) After the data has been sorted into customer account number order (separate program) the ledger update program is run.
- f) The computer prints out the total of invoices posted to the ledger and the control clerk reconciles this to the batch totals and then sends all the output documents to the sales department. The sales clerk again checks the total of invoices posted against his records.

4. DATA VET PROGRAM (DATA VALIDATION)

- a) Range Checks: these are designed to ensure that the data in a certain record field lies within predetermined limits e.g. rejection of clock card outside the range 10-80 hours. And print out special report for any clock card with hours over 80.
- b) Limit Checks:- similar to range checks but check that data is not below a certain value, or above a certain value.
- c) Existence Checks:- these are checks on record fields to ensure that the data is valid for that field.
e.g. - check record type is 1, 2, or 3.
- check that stock number exists.
- d) Format Checks: the record is checked to ensure that it has the required data fields and each data field is checked to ensure that the format (and size) of the data in it is correct e.g.

- numeric	N	N	N	N
- alphabetic	A	A	A	A
- alphc-numeric	A	N	N	N
- e) Sequence Checks: checking that records and batches are processed in correct sequence. Input in order e.g. cards.
- f) Completeness Checks: batch controls are reconciled to ensure that all records and batches have been processed.

- g) FIELD Comparisons and cross-checks - checks can confirm that the data in one field of the record is consistent with data in another field (e.g. if the sales area is 2, the customer number must be in the range 2000-2999).
- h) Check digit verification - check digits are added to key code numbers and used as an invaluable validation check to detect transposition and transcription errors in writing and keying in code numbers.

5. WHEN VALIDATION CHECK IDENTIFIES AN ERROR

- a) The record concerned will probably be rejected and processed no further. Rejection reports may be printed out during the data vet program, or they may be carried through the remaining programs and output with other printed data. All errors must be logged and corrected (usually by the user department) and re-input.
- b) The record concerned may not, on occasion be rejected, but an EXCEPTION REPORT (error report) output. This may happen, for example, where a range limit check is not satisfied. Just because a record is not within pre-set limits does not automatically mean the record is incorrect.
- c) In batch controls, the whole batch must be rejected for checking. For this reason the number of records in a batch should be kept reasonably low, say upto 30.

6. CHECH DIGIT VERIFICATION AS A DATA VET CHECK

It is very difficult for the human eye to detect error in a code on a casual inspection. Coding errors may escape detection before input to a computer system. To avoid the situation where the computer treats an incorrect code as the correct one, systems of self-checking numbers are used. The most common have check digits. These digits are added to the end of the code and give the whole new number some number special mathematical property. Check digits are commonly used in computer systems for key field identification codes. (e.g. customer number, employee number, student number).

6.1 THE CALCULATION OF A CHECK DIGIT (Using the common Modules 11 system)

example: Basic code number 5193.

	<u>DIGIT</u>	<u>WEIGHT</u>	
a) assign a weight to each digit commencing with 2 for the digit on the right, 3 for the next etc.	3	x 2	= 6
	9	x 3	= 27
	1	x 4	= 4
multiply each digit by its weight	5	x 5	= 25
b) Add up the results			<u>62</u>

- c) divide the total by the modulus 11 $62 \div 11 = 5$ remainder 7
- d) subtract the remainder from 11 $11 - 7 = 4$
- e) The result becomes the check digit, and the code number now becomes 51934. If the check digit is 10, the letter x is used.

6.2: The check digit is a number which, when weighted by 1 and added to the total in (a) above, will give a total, which is exactly divisible by 11.

$$4 \times 1 = 4$$

$$3 \times 2 = 6$$

$$9 \times 3 = 27$$

$$1 \times 4 = 4$$

$$5 \times 5 = 25$$

66

When the code is input to the computer, the data vet program will perform this calculation and if the result is not divisible by 11, the record will be rejected and a report output.

The Main Stages in Computer Based Sales Ledger

	<u>Processing Steps and Control</u>
SALES INVOICES	<u>Manual Preparation</u>
	1. Prepare invoices from order details despatch notes etc.
DATA CAPTURE	2. Check invoice calculations etc.
	3. Batch invoices, and set up batch controls
DATA PREPARATION	<u>DATA Preparation</u>
	1. Key in invoice and control data (onto punched cards- magnetic - tape etc)
INVOICE (AND CONTROL DETAILS) e.g. Punched Cards	2. Verify keying-in
	3. Correct errors
Computer Programs	<u>VALIDATION RUN</u>
DATA VET (Validation)	1. Validation checks made - rejection and exceptions reported
	2. Batch controls, rechecked, reconciled and provide total file controls.
	3. File of validated data prepared
SORT	<u>SORT</u>
MASTER FILE UPDATE	1. Invoices sorted into customer code sequence
	2. File controls checked and written onto new file of valid sorted data
OUTPUT ROUTINES	<u>UPDATE</u>
	1. B/f customer file records updated
	2. File controls
	- existence of supplies records
	- consistency with previous Sales,
DESPATCH AND CHECKING OUTPUT	<u>OUTPUT</u>
	1. Write control totals to c/f customer file
	2. Print control details

3. Print exceptional balance (e.g. Credit over limit)
4. Ensure reports etc sent to the right user department at right time.
5. Carry out manual reconciliation of controls after adjusting for rejected items.
