Chapter
2

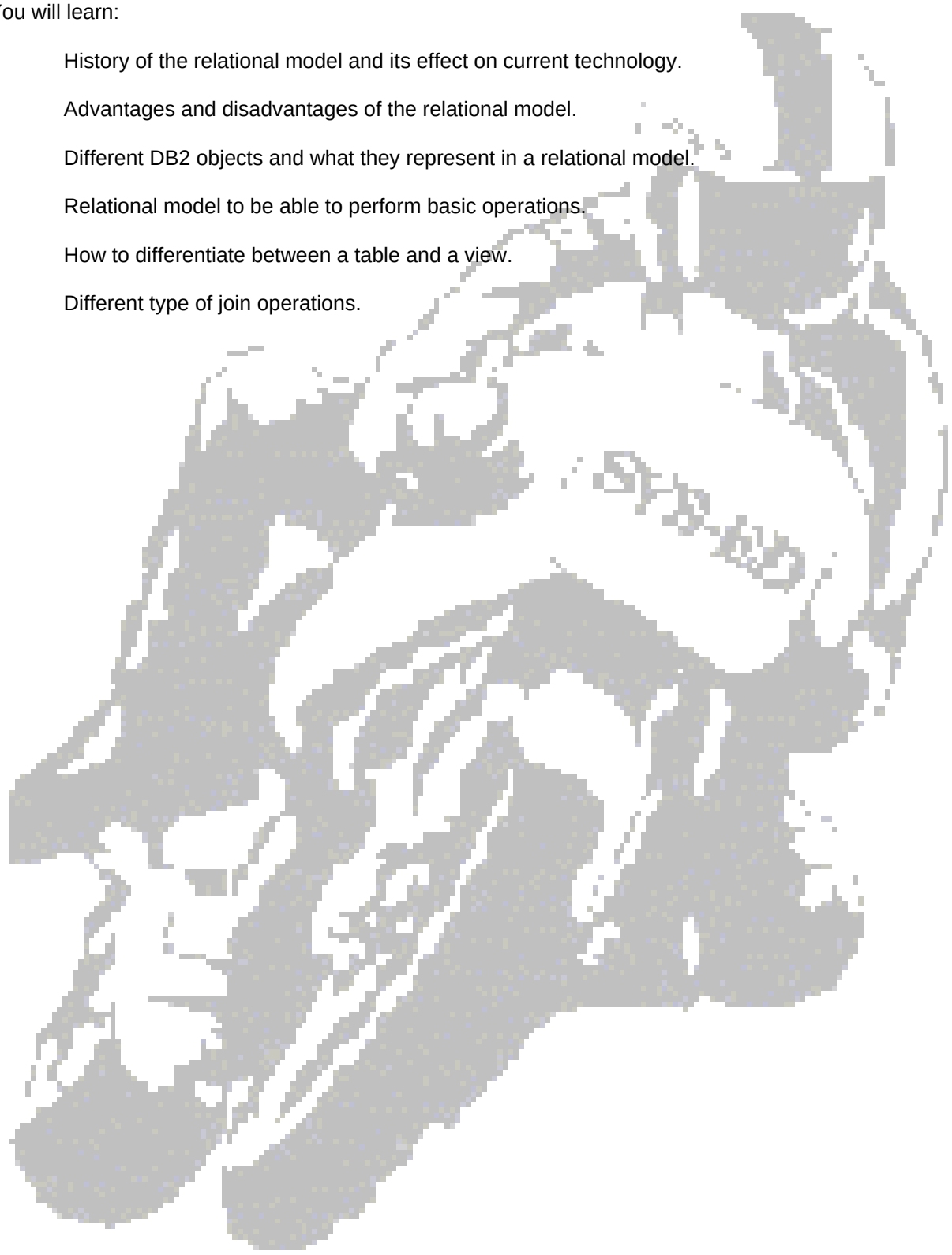
THE RELATIONAL MODEL

SYS-ED/
COMPUTER
EDUCATION
TECHNIQUES, INC.

Objectives:

You will learn:

- History of the relational model and its effect on current technology.
- Advantages and disadvantages of the relational model.
- Different DB2 objects and what they represent in a relational model.
- Relational model to be able to perform basic operations.
- How to differentiate between a table and a view.
- Different type of join operations.



1 History of Relational Model

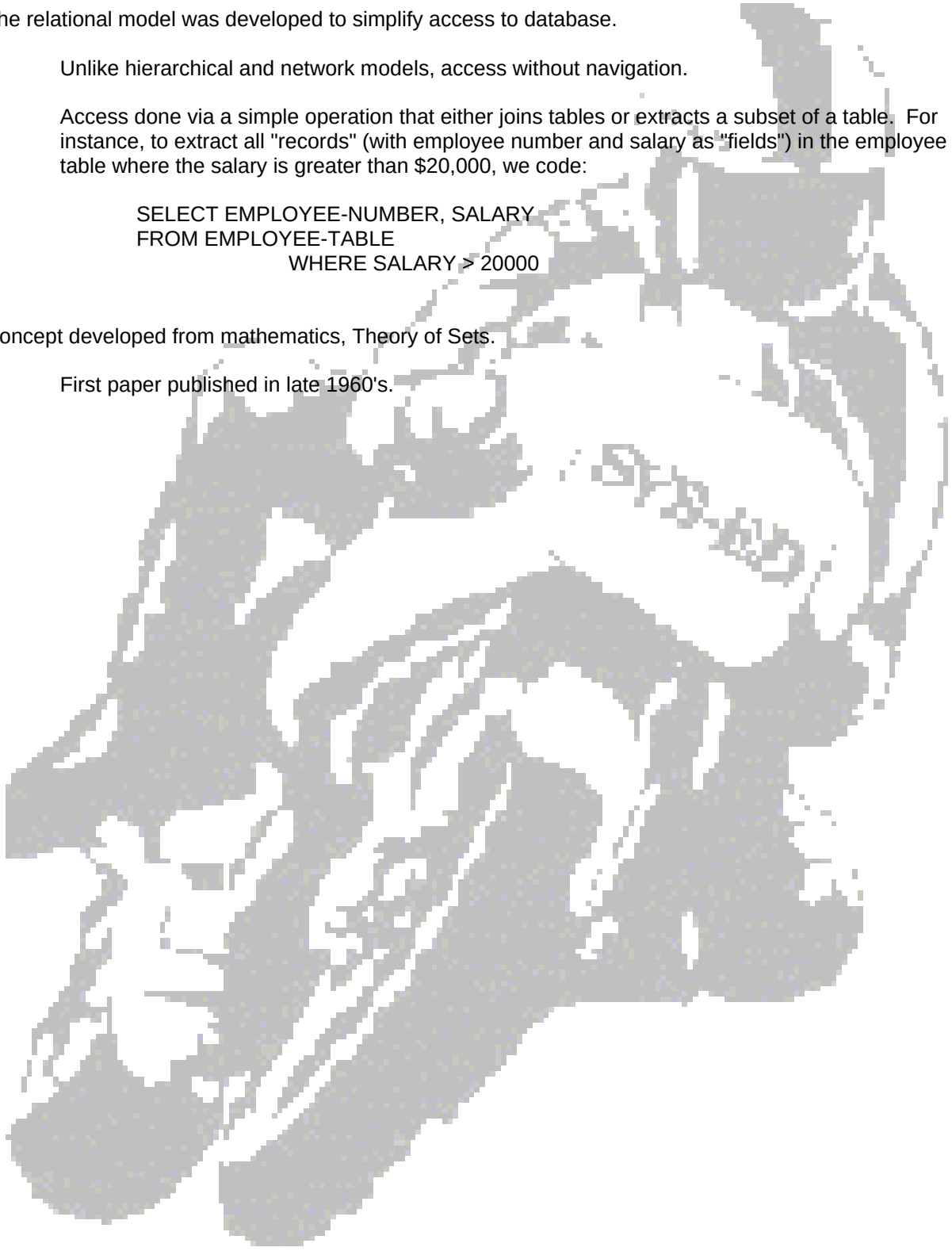
The relational model was developed to simplify access to database.

- Unlike hierarchical and network models, access without navigation.
- Access done via a simple operation that either joins tables or extracts a subset of a table. For instance, to extract all "records" (with employee number and salary as "fields") in the employee table where the salary is greater than $20,000, we code:

```
SELECT EMPLOYEE-NUMBER, SALARY
FROM EMPLOYEE-TABLE
        WHERE SALARY > 20000
```

Concept developed from mathematics, Theory of Sets.

- First paper published in late 1960's.



2 Relational Model of Database

Data in tables which are similar to sequential files.

- Row ====> "record"
- Column ====> "field"

Each table is independent.

- Programmer may use (when authorized) the table(s) needed.
- No parent/child linkage.

2.1 Characteristics of the Relational Model

Data exists in tables.

- Two dimensions only: row and column.

Each row represents one occurrence of an entity-type: one employee, one inventory item, etc.

- Row usually identified by unique key.
  Key is not mandatory in DB2.

Each column represents a property of an entity-type: salary, unit price, etc.

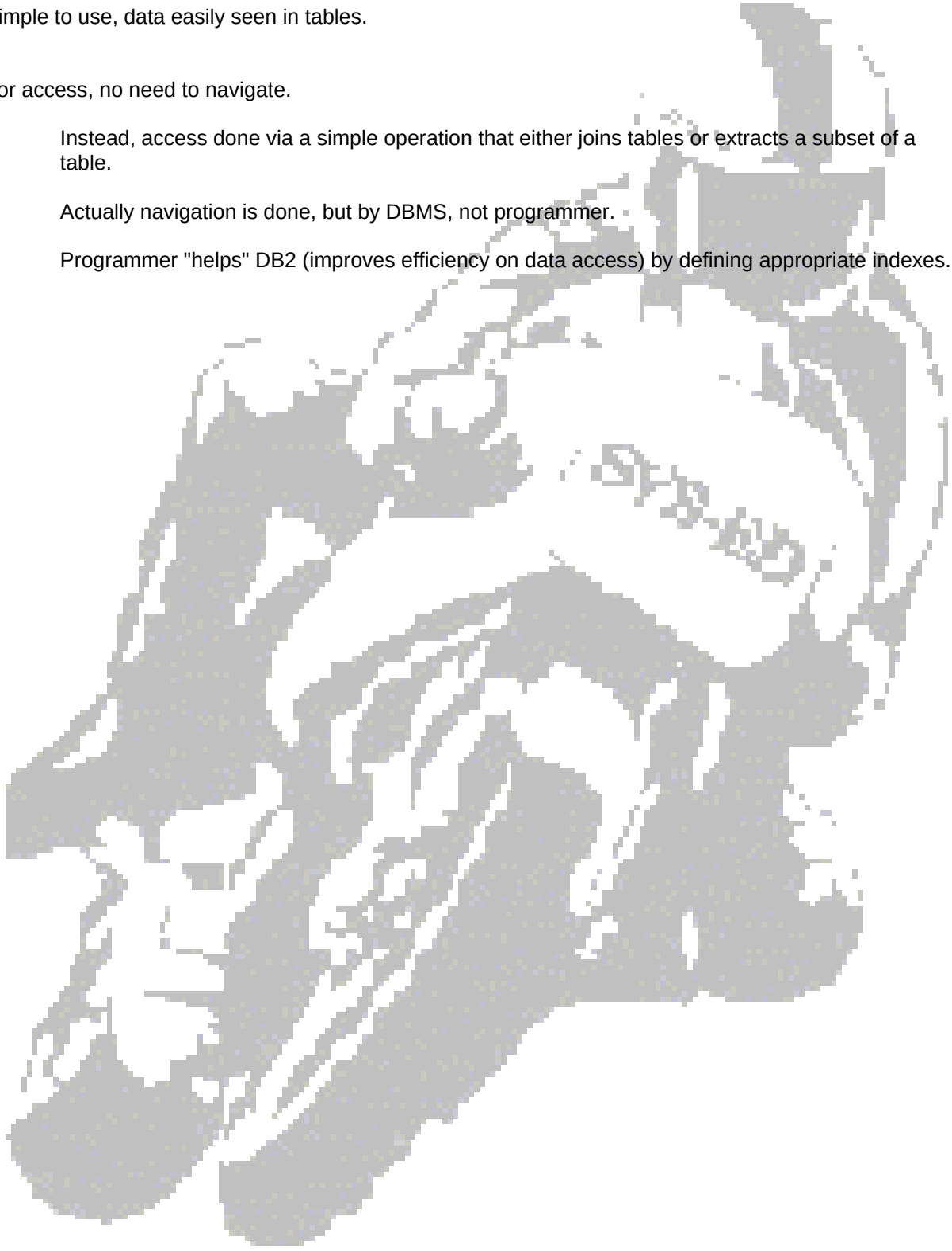
Table(s) may be manipulated to form new tables.

2.2 Advantages of the Relational Model

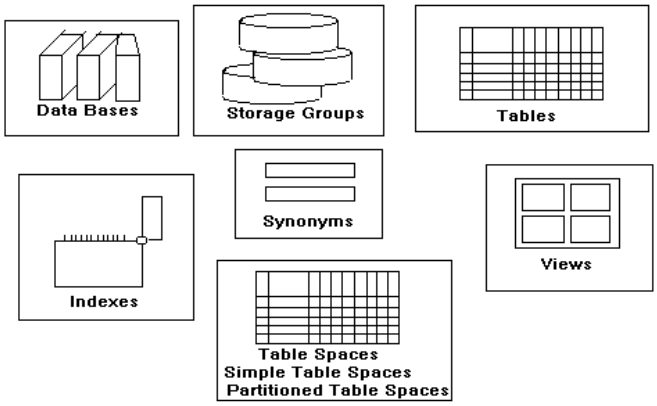
Simple to use, data easily seen in tables.

For access, no need to navigate.

- Instead, access done via a simple operation that either joins tables or extracts a subset of a table.
- Actually navigation is done, but by DBMS, not programmer.
- Programmer "helps" DB2 (improves efficiency on data access) by defining appropriate indexes.



3 DB2 Objects



A storage group is a set of DASD volumes of the same type.

- DB2 objects are physically contained in one or more storage groups.

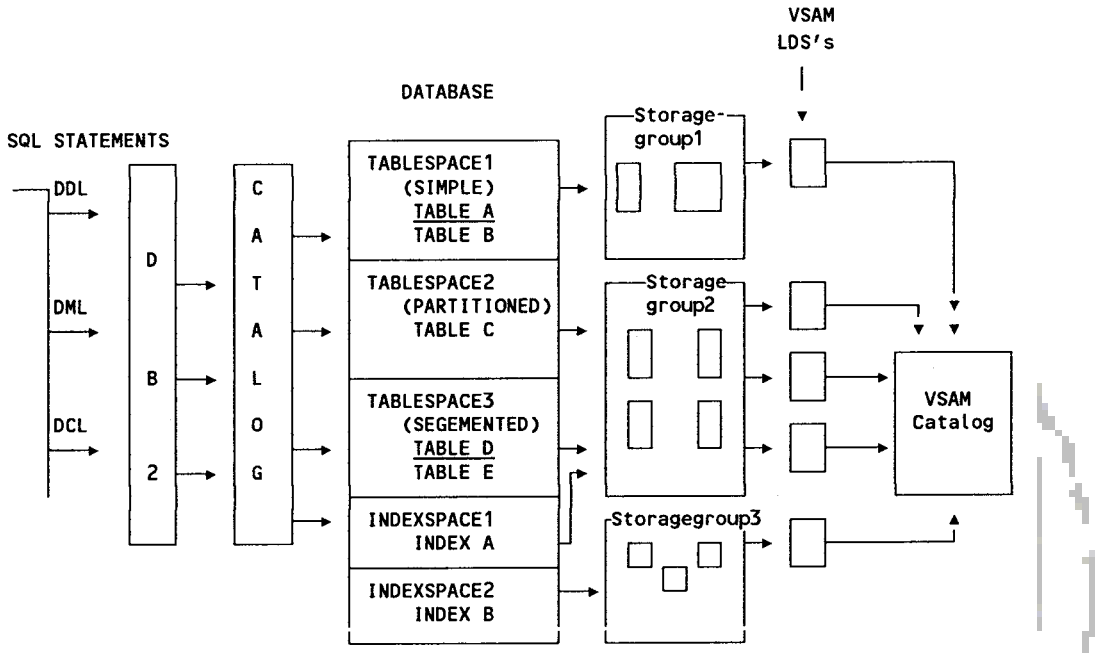
A database is a set of DB2 objects, usually logically related.

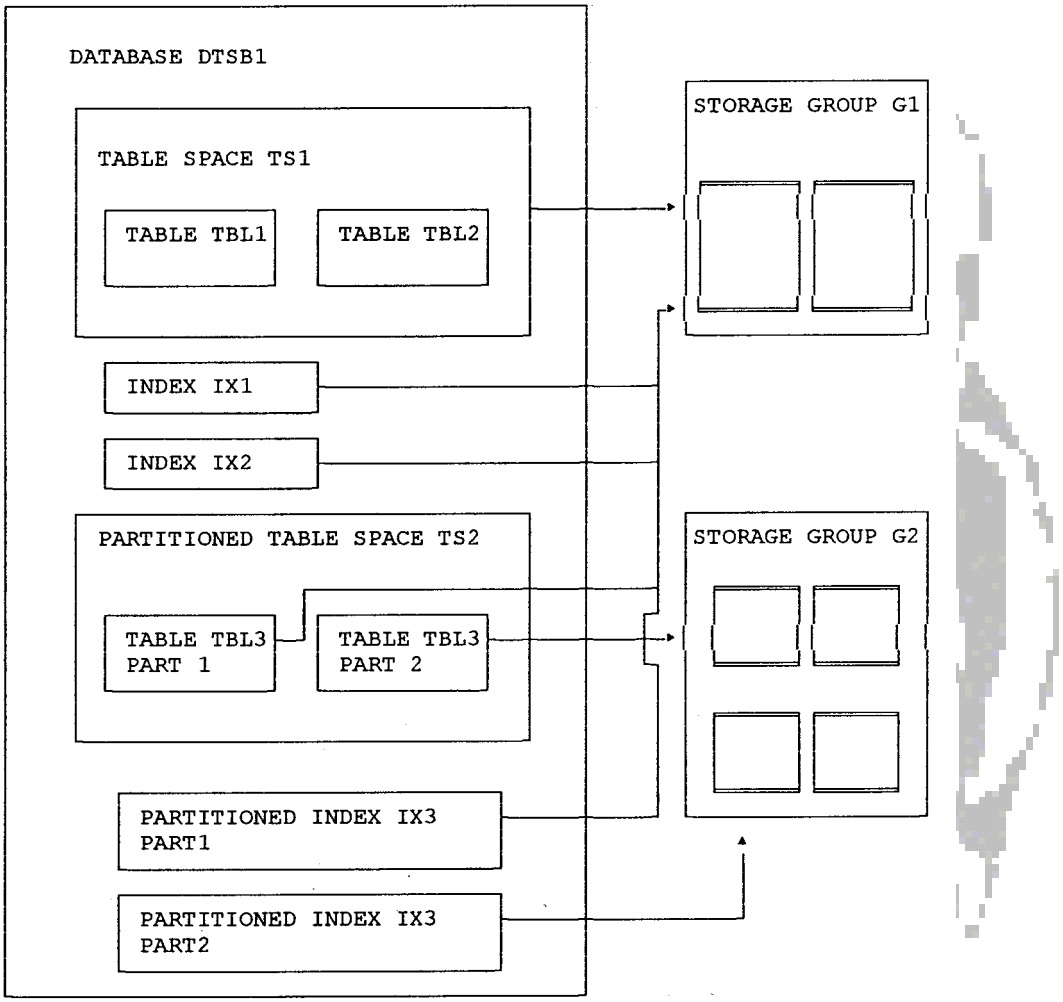
- For example, the employee table and the associated department table, as well as their indexes.

A tablespace is a set of VSAM linear files that contain tables.

- Many tables wholly contained in one tablespace.
- Or, one table only, but large tables are stored in one or more storage groups.
- Note: that tables are implemented as VSAM linear dataset or files, but transparent to user.

A table contains the rows of data, each row having one or more columns.


VSAM
LDS's
DATABASE
SQL STATEMENTS
DDL
DML
DCL
D
B
2
C
A
T
A
L
O
G
TABLESPACE1
(SIMPLE)
TABLE A
TABLE B
TABLESPACE2
(PARTITIONED)
TABLE C
TABLESPACE3
(SEGEMENTED)
TABLE D
TABLE E
INDEXSPACE1
INDEX A
INDEXSPACE2
INDEX B
Storage-
group1
Storage
group2
Storagegroup3
VSAM
Catalog


DATABASE DTSB1
TABLE SPACE TS1
TABLE TBL1
TABLE TBL2
INDEX IX1
INDEX IX2
PARTITIONED TABLE SPACE TS2
TABLE TBL3
PART 1
TABLE TBL3
PART 2
PARTITIONED INDEX IX3
PART1
PARTITIONED INDEX IX3
PART2
STORAGE GROUP G1
STORAGE GROUP G2

4 Creating DB2 Objects

Done by DBA, or systems administrator, not application programmer, using SQL DDL statements.

CREATE STOGROUP storage-group-name

- Creates storage group of specified volumes.
- Or use SYSDEDEFLT, the default storage group.

CREATE DATA BASE database-name

- Creates database in specified storage group.

CREATE TABLESPACE tablespace-name

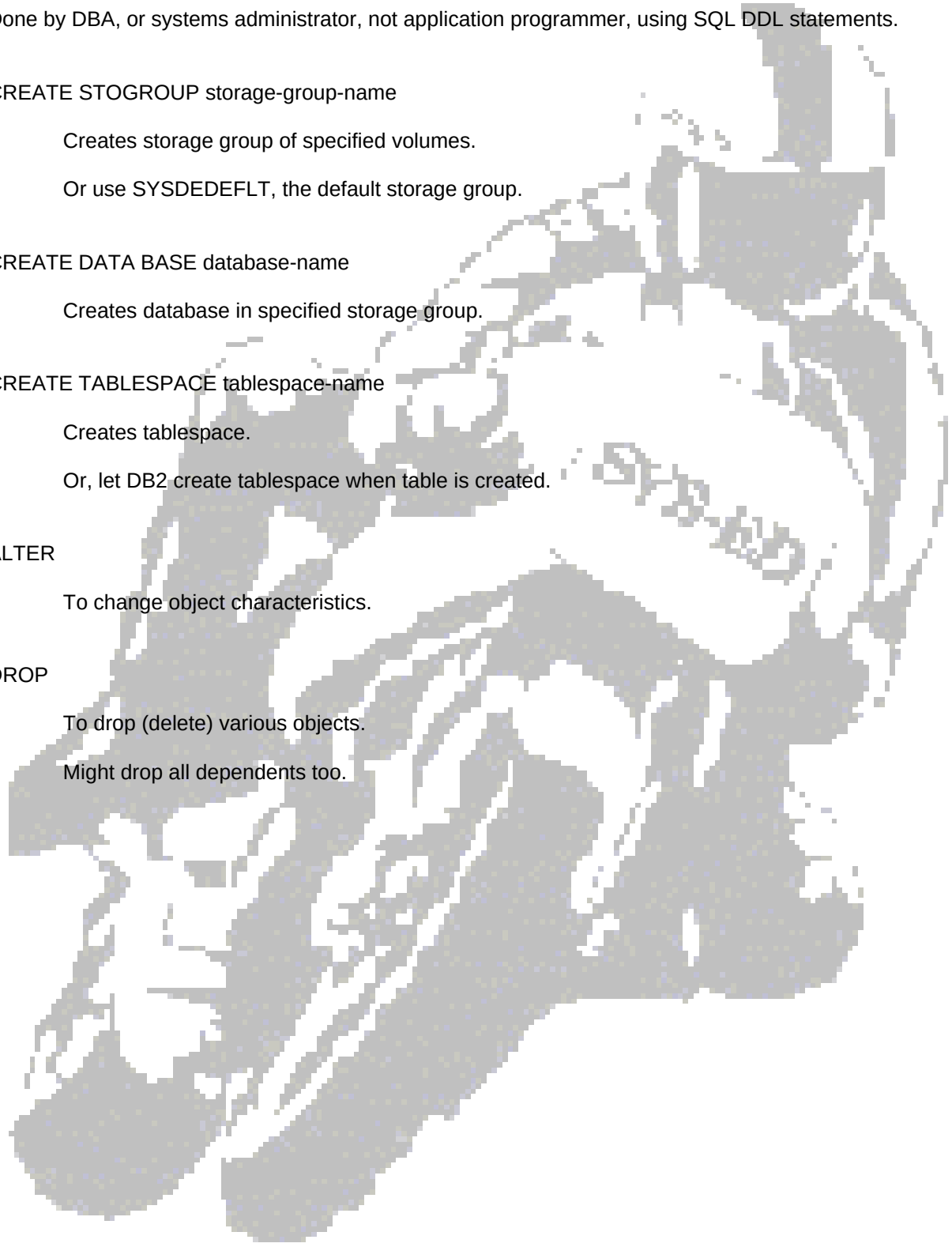
- Creates tablespace.
- Or, let DB2 create tablespace when table is created.

ALTER

- To change object characteristics.

DROP

- To drop (delete) various objects.
- Might drop all dependents too.



4.1 CREATE TABLE

```
CREATE TABLE tablename
        (column-name1 datatype (length)
              NULL|NOT NULL [WITH DEFAULT],
         column-name2.....)
        IN databasename.tablespacename
```

The most common datatypes/length are:

CHAR(n)	Character of n bytes length.
VARCHAR(n)	Variable-length character of n bytes maximum; saves DISK space and increases processing.
SMALLINT	Binary halfword.
INTEGER	Binary fullword.
DECIMAL(n,p)	Packed decimal of n digits length, p digits precision.

- Date
- Time
- Timestamp

NULL (the default) means a row is accepted on entry even without data on that column or is set to null.

NOT NULL means the column must contain data before the row is accepted and cannot be set to NULL. This is mandatory for primary key.

NOT NULL WITH DEFAULT means if there is no data on that column, DB2 generates data on that column, DB2 generates spaces for character, zeros for numeric, data types. It generates current date, current time, current timestamp for datatype date, time and timestamp.

5 Structure of a Table

Only two dimensions in table, rows and columns.

- Such a table is in first normal form; this normal form is mandatory for relational data base.
- There are other normal forms, all optional and having a bearing on efficiency.

If a table has repeating group (thus not of first normal form), it must be transformed.

- Repeating groups eliminated by moving them to other tables, or by having additional rows.

6 Unnormalized and Normalized

VENDOR NUMBER	ITEM	QTY	ITEM	QTY
001	0002	100	0003	30
002	0002	200		
003	0001	50	0004	100

Unnormalized

VENDOR NUMBER	ITEM NUMBER	QTY
001	0002	100
001	0003	30
002	0002	200
003	0001	50
003	0004	100

Normalized

7 Indexes

User may define one or more indexes for a table.

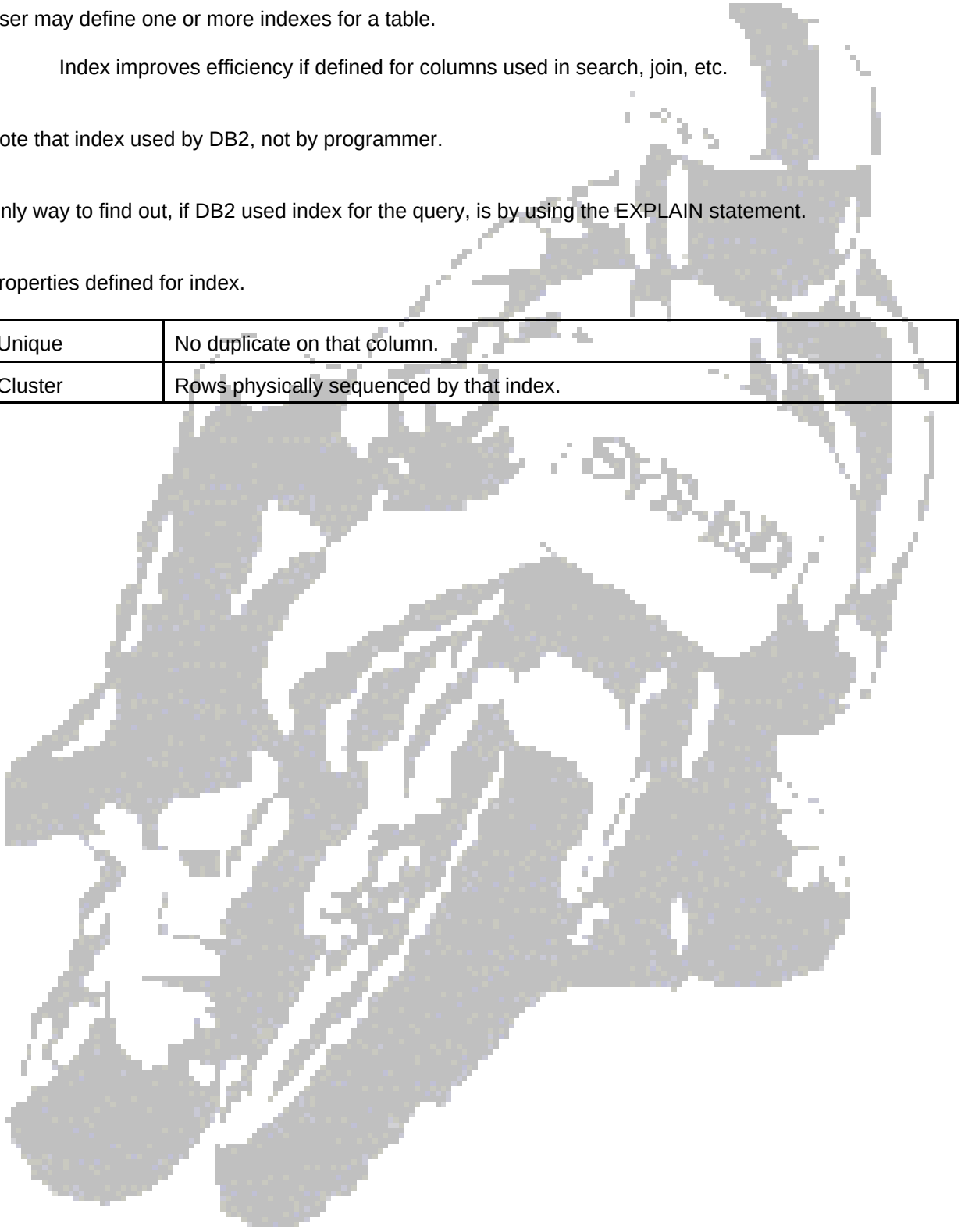
- Index improves efficiency if defined for columns used in search, join, etc.

Note that index used by DB2, not by programmer.

Only way to find out, if DB2 used index for the query, is by using the EXPLAIN statement.

Properties defined for index.

Unique	No duplicate on that column.
Cluster	Rows physically sequenced by that index.



8 Selection Operation

Gets horizontal subset of table (rows).

- In formal mathematics, all columns are retained.
- In DB2, SELECT (with WHERE clause): the columns may optionally be selected.

8.1 Select Statement

SELECT	list of table cols.
FROM	table/s and/or views.
WHERE	predicates for now selection.
GROUP BY	some of the selected cols.
HAVING	predicate for selecting groups.
ORDER BY	sorts the result table.

9 Purchase Order Relation

VENDOR NUMBER	ITEM NUMBER	QTY	ORDER DATE
001	0002	100	9/12/86
001	0003	30	9/12/86
002	0002	200	9/12/86
003	0002	50	9/13/86
003	0004	100	9/13/86

```
*
*
Select
*
*
?
```

VENDOR NUMBER	ITEM NUMBER	QTY
001	0002	100
002	0002	200

10 Projection Operation

Gets vertical subject of table (columns).

- DB2 SELECT (without WHERE clause).

10.1 Purchase Order Table

VENDOR NUMBER	ITEM NUMBER	QTY
001	0002	150
001	0003	30
002	0002	200
003	0001	50
003	0004	100

*
*
Project
*
*
?

ITEM NUMBER	QTY
0002	150
0003	30
0002	200
0001	50
0004	100

11 Joining Operation

Concatenate two or more tables.

- DB2 SELECT (with FROM clause).

JOIN OPERATION

Vendor Number

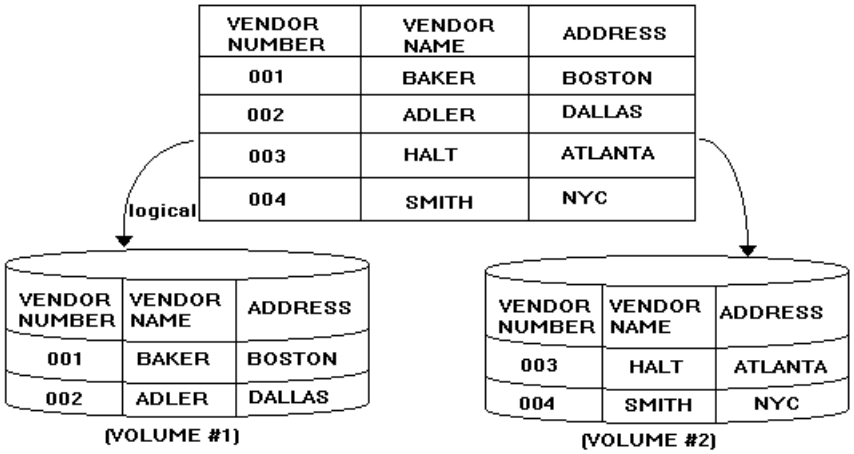
Vendor Number	Vendor Name	Address
001	BAKER	BOSTON
002	ADLER	SEATTLE
003	HOLT	ATLANTA

Purchase Order Table

Vendor Number	Item Number	QTY
001	0002	100
001	0003	30
002	0002	200
003	0001	50
003	0004	100

Vendor Number	Vendor Name	Item Number	QTY
001	BAKER	0002	100
001	BAKER	0003	30
001	BAKER	0002	200
001	BAKER	0001	50
001	BAKER	0004	100

11.1 Logical Tables Implemented on Physical Tables



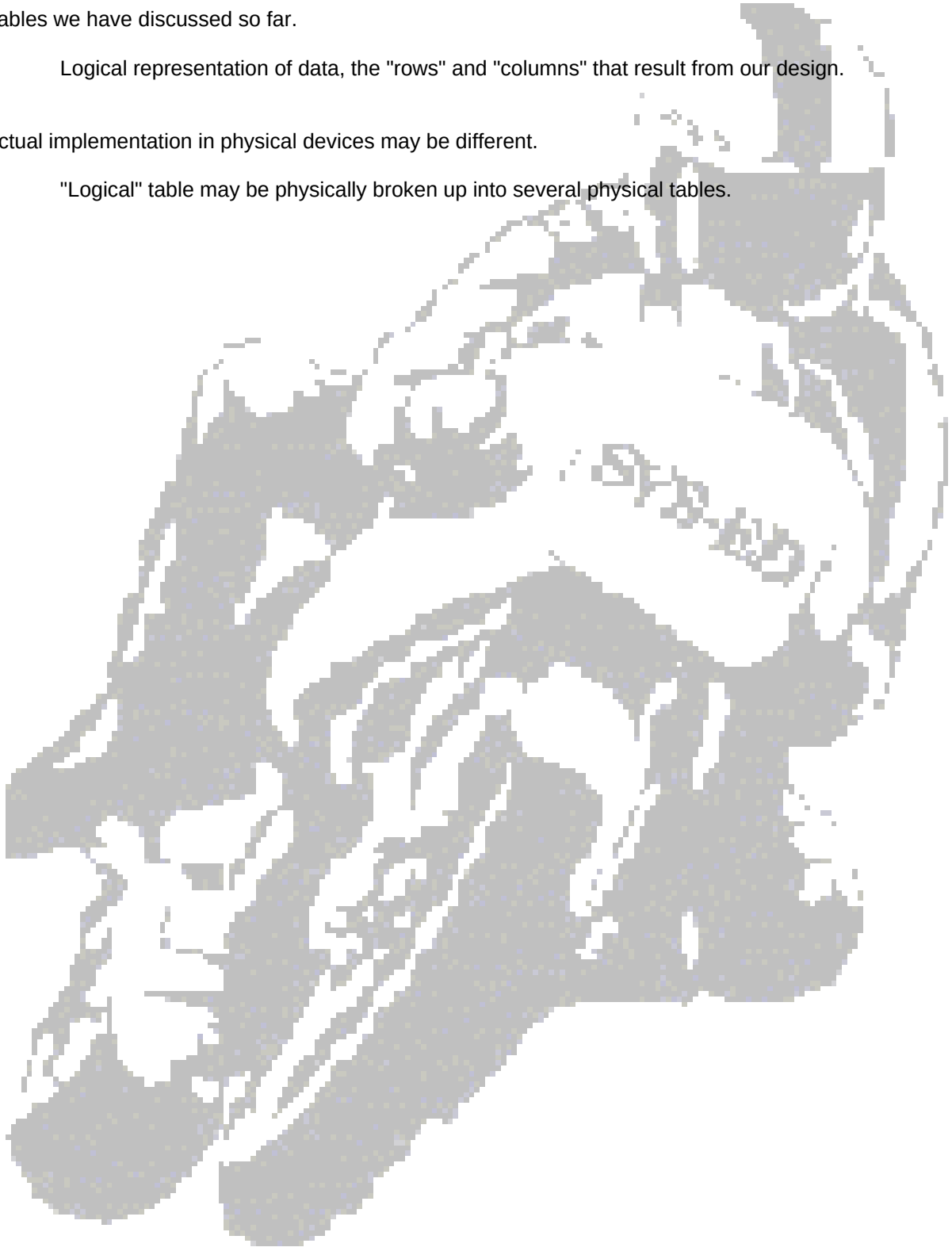
12 Real Tables

Tables we have discussed so far.

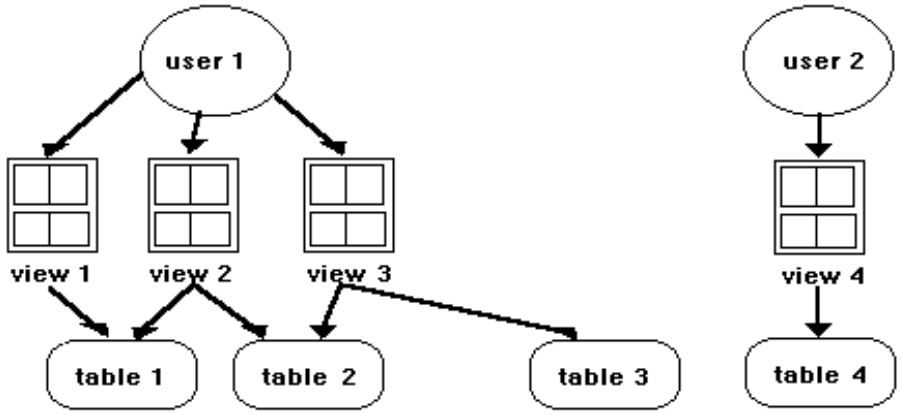
- Logical representation of data, the "rows" and "columns" that result from our design.

Actual implementation in physical devices may be different.

- "Logical" table may be physically broken up into several physical tables.



13 Views as Virtual Tables



A view is what a specific user is allowed to see (his/her own picture of the database).

- Each user has one or more views.

It does not have any physical data of its own, but just a definition in the DB2 catalog.

- Formed from one or more "real" tables.
- However, a user perceives it as a "real" table and except for some restrictions, uses the same operations used with "real" tables.
- Example of restriction.

  Update to view allowed only if view comes from one table. View resulting from a join or multiple tables cannot be updated.

Uses:

- Presents a simplified version of table(s) to a user.
- Security. Certain columns and rows not shown to user.