2021-DSE
ICT
PAPER 2A

HONG KONG EXAMINATIONS AND ASSESSMENT AUTHORITY
HONG KONG DIPLOMA OF SECONDARY EDUCATION EXAMINATION 2021

Please stick the barcode label here.

Candidate Number									

INFORMATION AND COMMUNICATION TECHNOLOGY

PAPER 2A

Databases

Question-Answer Book

11:15 am – 12:45 pm (1 hour 30 minutes)
This paper must be answered in English

INSTRUCTIONS

(1) After the announcement of the start of the examination, you should first write your Candidate Number in the space provided on Page 1 and stick barcode labels in the spaces provided on Pages 1, 3, 5 and 7.

(2) Answer **THREE** out of four questions. Write your answers in the spaces provided in this Question-Answer book. Do not write in the margins. Answers written in the margins will not be marked.

(3) Supplementary answer sheets will be supplied on request. Write your candidate number, mark the question number box and stick a barcode label on each sheet, and fasten them with string **INSIDE** this book.

(4) No extra time will be given to candidates for sticking on the barcode labels or filling in the question number boxes after the 'Time is up' announcement.

(5) The last page of this Question-Answer book contains SQL commands and symbols used in entity-relationship diagrams which you may find useful.

* A 2 0 0 E 0 2 A *

Answer THREE questions only.

1. An organisation holds an inter-school programming competition annually. It uses database tables SCHOOL, PSTUD and MARK to store information on schools, participating students and marks of students respectively.

The competition has two groups, the Junior group and Senior group. Participating students will be awarded a mark. There will be no record in MARK for a student who is absent from the competition.

SCHOOL

Field name	Type	Description	Example
SID	Character	School code	S0013
SNAME	Character	Name of school	ABC Government Secondary School

PSTUD

Field name	Type	Description	Example
SID	Character	School code	S0013
PID	Character	Identity code of student	P2020023
PNAME	Character	Name of student	Chan Siu Man
FIRST	Boolean	The student enrols in the competition for the first time	TRUE
GP	Character	Junior group (J) or Senior group (S)	J

MARK

Field name	Type	Description	Example
PID	Character	Identity code of student	P2020023
SMARK	Integer	Mark awarded in the competition	74

Write SQL statements to complete the following tasks from (a) to (c) below.

(a) List the names of the students who come from the school with the school code 'S0013'.

(2 marks)

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

Please stick the barcode label here.

(b) Find the total number of students who come from the schools with names including 'Government'.

(3 marks)

(c) List the names of students in the Junior group who get a mark greater than or equal to 60.

(3 marks)

(d) What is the purpose of the following SQL statement?

```
SELECT PID, PNAME FROM PSTUD
WHERE GP = 'S' AND FIRST
  AND PID NOT IN (SELECT PID FROM MARK)
```

(2 marks)

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

(e) Complete the following SQL statement for finding the sum of the marks of the students in each school.

```
SELECT S.SID, _____ AS TOTAL
FROM  SCHOOL S, PSTUD P, MARK M
WHERE _____
      _____
GROUP BY _____
```

(3 marks)

Database table TM stores the results of the SQL statement in (e).

TM

Field name	Type	Description
SID	Character	School code
TOTAL	Integer	The sum of the marks of students

(f) A school will be awarded a certificate if the sum of the marks of its students is the highest among the schools. Write a SQL statement with TM to find the name(s) of school(s) that will be awarded a certificate.

(2 marks)

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

Please stick the barcode label here.

2. A library uses database tables READER, BOOK, CIR and BKCOPY to store information on readers, books, circulation and copies of books respectively.

READER

Field name	Description	Example
RID	Identity code of reader	R0132
NAME	Name	Chan Tai Man

BOOK

Field name	Description	Example
BID	Identity code of book	B102
TITLE	Title	A Brief History of Time
CAT	Category	Science
AUTHOR	Author	Stephen Hawking

CIR

Field name	Description	Example
ITEMNO	Item number	B102C1
RID	Identity code of reader	R0132
DOB	Timestamp of borrowing	1/12/2020 10:20
DOR	Timestamp of return	20/12/2020 15:30
FINE	Overdue fine ($5 per day)	25

BKCOPY

Field name	Description	Example
ITEMNO	Item number	B102C1
BID	Identity code of book	B102
DOP	Date of purchase	15/10/1990

Books in the library can be borrowed for a period of 14 days. There may be several copies of certain books and each copy has a unique item number ITEMNO stored in BKCOPY.

(a) Identify **two** candidate keys of CIR.

(1) _____

(2) _____

(2 marks)

(b) (i) Which field can be regarded as a derived attribute? Explain briefly.

_____

_____

_____

(2 marks)

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

(ii) State a reason for having a derived attribute in a database.

_____

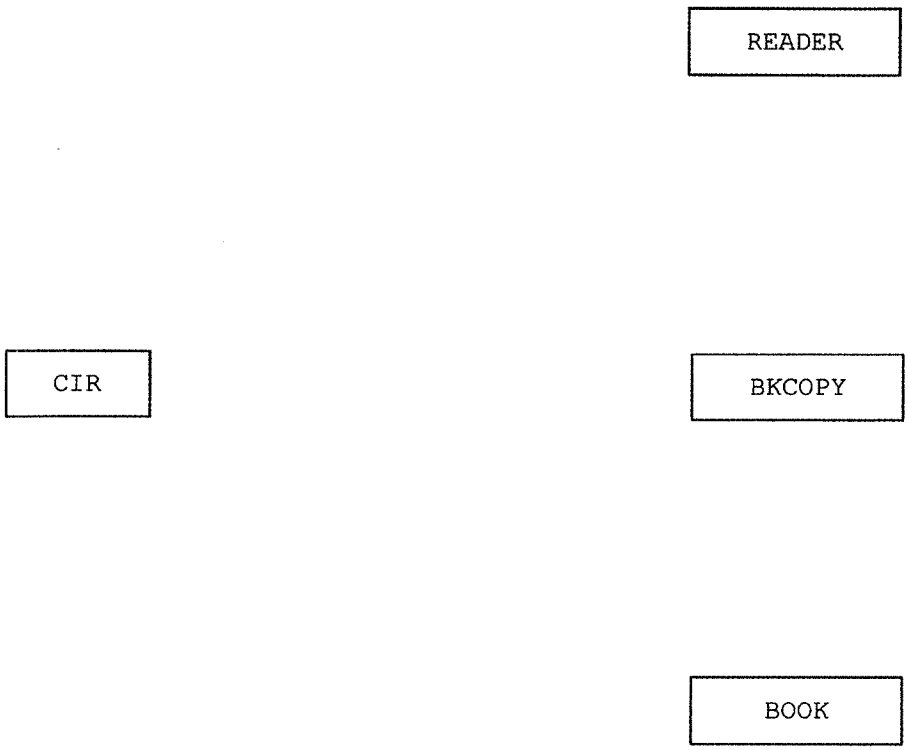
_____

(1 mark)

(c) The requirements of the database are described below:

Each reader may borrow a maximum of 5 books. Some readers may not have borrowed any books from the library. Some books may never be borrowed. There is more than one copy of some books.

Complete the ER diagram below for this database. It is not necessary to draw attributes.



(6 marks)

Answers written in the margins will not be marked.

Please stick the barcode label here.

Part of the records in the database are as follows:

READER

RID	NAME
R0132	Chan Tai Man
R0124	Lee Ka Ka
R0155	Lee Ka Ka

Primary key: RID

CIR

ITEMNO	RID	DOB	DOR	FINE
B102C1	R0132	1/12/2020	20/12/2020	25
B134C1	R0132	20/12/2020		
B134C2	R0124	11/12/2020	18/12/2020	0

(d) Explain the integrity problem in the database for each of the following cases.

(i) RID in CIR is replaced by NAME.

(2 marks)

(ii) The record of the reader 'Chan Tai Man' in READER is deleted.

(2 marks)

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

3. John and Mary organise a singing contest in a school. Students join the contest in teams. The database tables SONG and TEAM are used to store information on songs and teams respectively.

SONG

Field name	Description	Example
SID	Identity code of song	0117
TITLE	Song title	Happy birthday
DS	Duration of song	6:15

Primary key: SID

TEAM

Field name	Description	Example
TID	Identity code of team	024
TNAME	Team name	Rainbow
SEQ	Order of performance	8
TMARK	Mark awarded in the contest	78

Primary key: TID

(a) (i) Complete the following SQL statement for preventing the same order of performance from entering in SEQ.

```
CREATE TABLE TEAM (TID char(3) primary key,
                   TNAME char(30),
                   SEQ int _____,
                   TMARK int NOT NULL)
```

(1 mark)

(ii) What is the pros and cons of using NOT NULL constraint on TMARK in the SQL statement above?

_____

_____

_____

_____

(2 marks)

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

(b) SONG and TEAM contain some records. John and Mary propose two methods that both use SID to link the two database tables together.

Method 1: Dropping TEAM and then re-creating this table with SID.
Method 2: Changing the structure of TEAM with SQL statements.

(i) What are the consequences of using Method 1?

_____

_____

_____

_____

(2 marks)

(ii) Complete the following SQL statements for Method 2.

```
ALTER TABLE TEAM
    ADD column SID char(4)

ALTER TABLE TEAM
    ADD _____ key (SID) _____ SONG(SID)
```

(2 marks)

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

STUDENT is a database table that stores the information on students.

STUDENT

Field name	Description	Example
STUDID	Identity code of student	2018103
NAME	Student name	Chan Ka Yan
TEL	Phone number	34567890

Primary key: STUDID

Students join the contest in teams of two to four members. There is only one team leader in each team. John and Mary propose two different methods to store the information on team leaders and team members.

(c) Mary proposes additional fields in TEAM, as shown below:

Field name	Description	Example
TID	Identity code of team	024
TNAME	Team name	Rainbow
SEQ	Order of performance	8
SID	Identity code of song	0117
TMARK	Mark awarded in the contest	78
LEADER	Identity code of student who is the team leader	2018103
MEM1	Identity code of student who is team member 1	2018112
MEM2	Identity code of student who is team member 2	2018120
MEM3	Identity code of student who is team member 3	

(i) To store team information, John proposes a new database table consisting of TID, an existing field and a new Boolean field. Complete the following design of the table proposed by John.

Field name	Description
TID	Identity code of team

(2 marks)

(ii) Give one advantage of John's proposal over Mary's proposal.

_____

_____

(1 mark)

(iii) Give one advantage of Mary's proposal over John's proposal.

_____

_____

(1 mark)

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

John wants to post a report about the results of the competition online for the public, as shown below:

Results of the competition

Order of Performance	Team	Song Identity Code	Song title	Student Identity Code	Student Name	Phone	Mark
1	Rainbow	0117	Happy birthday	2018103	Chan Ka Yan	34567890	78
1	Rainbow	0117	Happy birthday	2018112	Wong Ka Ming	23456789	78
1	Rainbow	0117	Happy birthday	2018120	Li Lai Kit	98765432	78
2	Thunder	0115	One day	2017138	Cheung Hoi Yan	22334455	80

⋮

Mary improves the report to meet the following requirements:

- To maintain data privacy
- To show the champion, first runners-up and second runners-up only
- To reduce redundant information
- To make the layout more readable

(d) Re-design the report and annotate your design, where appropriate.

(4 marks)

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

4. P and Q are two large chain stores. Recently P acquired Q. Their membership databases will be merged in one of the stages of the database application development lifecycle.

(a) Tim is responsible for most of the work in the requirements collection and analysis stage.

(i) What database personnel should Tim be? _____

(1 mark)

(ii) Describe **two** deliverables for this stage.

_____

_____

_____

(2 marks)

MEMP and MEMQ are database tables with the same field names that store information on members of P and Q respectively.

MEMP

Field name	Description	Example
MID	Identity code of member	K123456789
NAME	Member name	Wong Siu Mei
TEL	Phone number	98761234
GENDER	F = Female, M = Male	F

MEMQ

Field name	Description	Example
MID	Identity code of member	QQ456
NAME	Member name	Wong Siu Mei
TEL	Phone number	98761234
GENDER	0 = Female, 1 = Male	0

(b) According to MEMP and MEMQ, give **three** examples to explain why data conversion is necessary.

_____

_____

_____

_____

_____

_____

_____

_____

(3 marks)

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

(c) Tim considers the following two methods to merge two database tables.

Method 1: Use SQL statements to do the merging.
Method 2: Using a general programming language to write a program to do the merging.

(i) Give an advantage of Method 1 over Method 2.

(1 mark)

(ii) Give an advantage of Method 2 over Method 1.

(1 mark)

(d) Describe how data mining can be used to determine sales strategies.

(3 marks)

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

(e) Members' phone numbers are stored in the merged database.

(i) Phone number in the database is a candidate key instead of a primary key. Why?

(2 marks)

(ii) Give two potential benefits of having several candidate keys such as phone number in the database.

(2 marks)

END OF PAPER

Answers written in the margins will not be marked.

Answers written in the margins will not be marked.

Database (SQL commands - based on SQL-92 Standard)

Constants	FALSE, TRUE
Operators	+, –, *, /, >, <, =, >=, <=, <>, %, _ , ' , AND, NOT, OR
SQL	ABSOLUTE (ABS), AVG, INT, MAX, MIN, SUM, COUNT ASC, AT, CHAR (CHR), CHAR_LENGTH (LEN), LOWER, TRIM, SPACE, SUBSTRING (SUBSTR/MID), UPPER, VALUE (VAL) DATE, DAY, MONTH, YEAR ADD, ALL, ALTER, ANY, AS, ASC, BETWEEN, BY, CREATE, DELETE, DESC, DISTINCT, DROP, EXISTS, FROM, GROUP, HAVING, IN, INDEX, INNER JOIN, INSERT, INTEGER, INTERSECT, INTO, LEFT [OUTER] JOIN, LIKE, MINUS, NULL, RIGHT [OUTER] JOIN, FULL [OUTER] JOIN, ON, ORDER, SELECT, SET, TABLE, TO, UNION, UNIQUE, UPDATE, VALUES, VIEW, WHERE

Symbols Used in Entity-Relationship Diagrams

Meaning	Symbol	Meaning	Symbol
Entity	Entity	One-to-One Relationship	1 Relationship 1
Attribute	Attribute	One-to-Many Relationship	1 Relationship M
Key Attribute	Attribute	Many-to-Many Relationship	M Relationship N
Relationship	Relationship	Participation constraints: Use | on Mandatory side Use ○ on Optional side	| Relationship ○