CTT102 - Cơ sở dữ liêu

Tháng 1/2013

Truy vấn nâng cao

Tóm tắt nội dung bài thực hành:

Sử dụng một số hình thức truy vấn khác: Truy vấn phục vụ mục đích thống kê, ...

Bộ môn **Hệ thống thông tin**

Khoa Công nghệ thông tin

ĐH Khoa học tự nhiên TP HCM



MỤC LỤC

1 M	Mục tiêu và tóm tắt nội dung	1
2 H	Hướng dẫn chi tiết	1
2.1	Truy vấn với phép kết ngoài	\bigcirc 1
2.2	Tổng hợp dữ liệu sử dụng COMPUTE, COMPUTE BY, CUBE, ROLLUP	 5
2.3	.1	
2.4	Cấu trúc CASE	9
2.5	GROUP BY WITH CUBE ROLLUP với GROUPING	10
2.6		
3 B	Bài tập tại lớp	
	Bài tập về nhà	
	50, mon Hirit	

TRUY VẤN NÂNG CAO

1 Mục tiêu và tóm tắt nội dung

Sau khi hoàn thành bài thực hành này sinh viên sẽ biết được:

- Truy vấn với Cấu trúc CASE
- Sử dụng phép kết ngoài
- Tổng hợp dữ liệu sử dụng COMPUTE, COMPUTE BY, CUBE, ROLL UP
- Câu lệnh INSERT UPDATE mở rộng
- Sử dụng GROUP BY ... WITH CUBE | ROLLUP với hàm GROUPING

2 Hướng dẫn chi tiết

2.1 Truy vấn với phép kết ngoài

- Xét 2 bảng dữ liệu sau:

SINHVIEN				LOP			
	malop	tenlop			masv	hoten	malop
1	L1	10A		1	01	A	L1
2	L2	10B	•	2	02	В	L2
3	L3	10C		3	03	С	L2
				4	04	D	L1
				5	05	E	L1

- Yêu cầu: Cho biết sĩ số của mỗi lớp

Inner joins (Kết bằng):

Phép kết Inner joins giữa 2 bảng A và B → là một bảng C = {các bộ trong đó mỗi bộ là sự kết hợp của các bộ trong A với các bộ trong B sao cho điều kiện kết được thỏa mãn}

• Phép kết inner join giữa SINHVIEN và LOP

Sử dụng phép JOIN:

SELECT *

FROM SINHVIEN AS SV *JOIN* LOP AS L *ON* SV.MALOP = L.MALOP

Hoặc

Sử dụng điều kiện kết ở mệnh đề WHERE:

SELECT *

FROM SINHVIEN SV, LOP L

WHERE SV.MALOP = L.MALOP

<u>Kết quả</u>

	masv	hoten	malop	malop	tenlop
1	01	A	L1	L1	10A
2	02	В	L2	L2	10B
3	03	С	L2	L2	10B
4	04	D	L1	L1	10A
5	05	E	L1	L1	10A

Nhận xét: Thông tin về lớp 10C bị mất

• Tính sĩ số của lớp

SELECT L.MALOP, L.TENLOP, COUNT(*) AS SISO

FROM SINHVIEN AS SV JOIN LOP AS L ON SV.MALOP = L.MALOP

GROUP BY L.MALOP, L.TENLOP

Kết quả :

	malop	tenlop	SiSo
1	L1	10A	3
2	L2	10B	2

Nhận xét : Sĩ số của lớp 10C (bằng 0) không được xuất ra, vì thông tin lớp 10C đã bị mất sau phép kết bằng. Do vậy với phép kết bằng thì những bộ không thỏa mãn điều kiện kết sẽ được loại bỏ. Nếu muốn xuất hiện → Sử dụng phép kết ngoài.

Right (Outer) joins (Kết phải)

Phép kết Right Outer joins giữa 2 bảng A và B → là một bảng C = {các bộ trong đó mỗi bộ là sự kết hợp của các bộ trong A với các bộ trong B sao cho điều kiện kết được thỏa mãn} + {các bộ còn lại trong B mà không thỏa điều kiện kết với bất kỳ một bộ trong A nào}

Phép kết Right (Outer) Joins giữa SINHVIEN và LOP

SELECT *

FROM SINHVIEN as SV RIGHT JOIN LOP as L ON SV.MALOP = L.MALOP $K\acute{e}t$ $qu\dot{a}$:

	masv	hoten	malop	malop	tenlop
1	01	A	L1	L1	10A
2	04	D	L1	L1	10A
3	05	E	L1	L1	10A
4	02	В	L2	L2	10B
5	03	С	L2	L2	10B
6	NULL	NULL	NULL	L3	10C

Nhân xét: Thông tin về lớp 10C vẫn được giữ lại sau phép kết phải

• Tính sĩ số của lớp

SELECT L.MALOP, L.TENLOP, COUNT(SV.MALOP) AS SISO

FROM SINHVIEN AS SV RIGHT JOIN LOP AS L ON SV.MALOP = L.MALOP GROUP BY L.MALOP, L.TENLOP

Kết quả :

malop		tenlop	SiSo
1	L1	10A	3
2	L2	10B	2
3	L3	10C	0

<u>Nhận xét</u>: Sĩ số của các lớp không có học sinh (10 C) vẫn được xuất ra (vì phép kết không mất thông tin về lớp)

Left (Outer) joins (Kết trái)

Phép kết Left (Outer) joins giữa 2 bảng A và B → là một bảng C = {các bộ trong đó mỗi bộ là sự kết hợp của các bộ trong A với các bộ trong B sao cho điều kiện kết được thỏa mãn} + {các bộ còn lại trong A mà không thỏa điều kiện kết với một bộ bất kỳ trong B nào}

Full (Outer) joins

Phép kết Full Outer joins giữa 2 bảng A và B → là một bảng C = {các bộ trong đó mỗi bộ là sự kết hợp của các bộ trong A với các bộ trong B sao cho điều kiện kết

được thỏa mãn} + {các bộ còn lại trong A mà không thỏa điều kiện kết với bất kỳ một bộ trong B nào} + {các bộ còn lại trong B mà không thỏa điều kiện kết với bất kỳ một bộ trong A nào}

PUBLISHER

	pid	pname	pcity
1	1	Algodata Infosystems 1	MIA
2	2	Algodata Infosystems 2	NYO
3	3	Algodata Infosystems 3	MAN

<u>AUTHORS</u>

	auid	firstname	lastname	city
1	1	Reginald	Blotchet-Halls	NYO
2	2	Michel	DeFrance	OAS
3	3	Innes	del Castillo	CAN
4	4	Ann	Dull	LND
5	5	Marjorie	Green	CAL
6	6	Morningstar	Greene	CAL
7	7	Burt	Gringlesby	Los
8	8	Sheryl	Hunter	NYO

SELECT *

FROM AUTHORS **FULL JOIN** PUBLISHERS **ON** city = pcity

<u>Kết quả :</u>

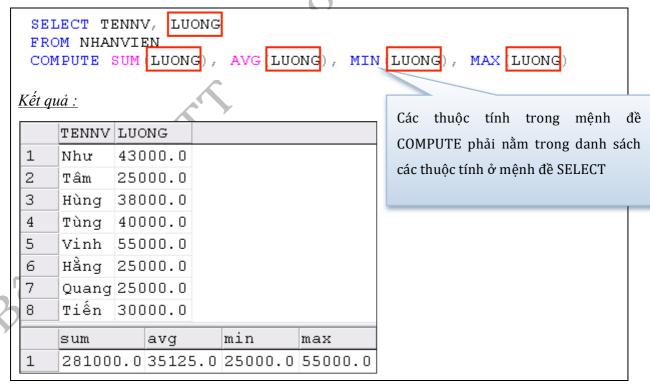
	auid	first	lastname	city	pid	pname	pcity	
1	1	Reginald	Blotch	NYO	2	Algoda	NYO	
2	2	Michel	DeFrance	OAS	NULL	NULL	NULL	
3	3	Innes	del Ca	CAN	NULL	NULL	NULL	
4	4	Ann	Dull	LND	NULL	NULL	NULL	
5	5	Marjorie	Green	CAL	NULL	NULL	NULL	
6	6	Morni	Greene	CAL	NULL	NULL	NULL	
7	7	Burt	Gringlesby	LOS	NULL	NULL	NULL	
8	8	Sheryl	Hunter	NYO	2	Algoda	NYO	
9	NULL	NULL	NULL	NULL	3	Algoda	MAN	
10	NULL	NULL	NULL	NULL	1	Algoda	MIA	

2.2 Tổng hợp dữ liệu sử dụng COMPUTE, COMPUTE BY, CUBE, ROLLUP

COMPUTE:

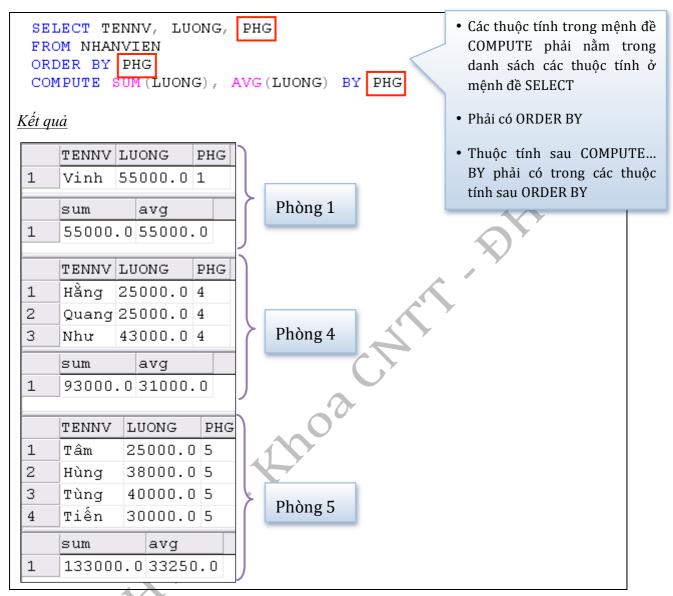
Sử dụng để tổng hợp dữ liệu của các bảng.

<u>Ví dụ 1</u>: Cho biết các nhân viên, tổng lương, lương trung bình của tất cả các nhân viên



COMPUTE BY

<u>Ví dụ 2</u>: Cho biết các nhân viên của từng phòng, tổng lương, lương trung bình của từng phòng



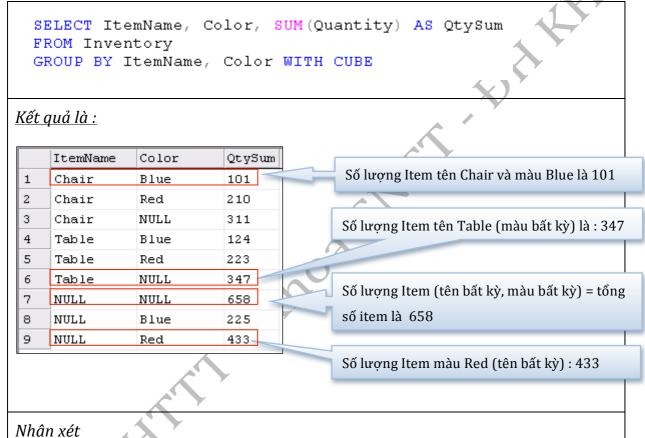
- Lưu ý :
 - Các thuộc tính sau COMPUTE ... BY phải có trong danh sách các thuộc tính sau ORDER BY
 - o Không đặt tên kết quả trả ra được

GROUP BY ... WITH CUBE

Cho bảng dữ liệu ITEM như sau:

	itemid	itemname	color	quantity
1	1	Table	Blue	124
2	2	Table	Red	223
3	3	Chair	Blue	101
4	4	Chair	Red	210

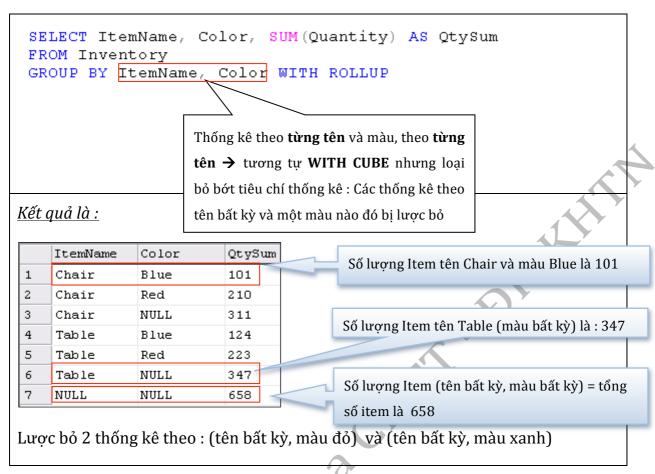
Ví dụ 3: Tổng hợp số lượng của các item theo tên và màu, theo từng tên, theo từng màu, tổng số item



- Group by n thuộc tính → sẽ thống kê theo 2ⁿ tiêu chí
- Những thống kê mà không có dữ liệu sẽ không được xuất ra

GROUP BY ... WITH ROLLUP

Ví du 4: Tổng hợp số lượng của các item theo từng tên và màu, theo từng tên, tổng số item



2.3 Các câu lệnh INSERT, UPDATE mở rộng

Cho các quan hệ sau:

SINHVIEN (MASV, HOTEN, DIEMTB, HANG)

SINHVIENGIOL (MASV, HOTEN, DIEMTB)

UPDATE dữ liệu từ dữ liệu có sẵn:

Ví dụ 5: Cập nhật hạng của sinh viên

```
update SINHVIEN
set HANG = (SELECT count(*)
FROM SINHVIEN sv
WHERE sv.DIEMTB >= SINHVIEN.DIEMTB)
```

INSERT dữ liệu vào một bảng từ một bảng có sẵn

Ví dụ 6: Thêm dữ liệu vào bảng SINHVIENGIOI các sinh viên có điểm trung bình từ 8.0 trở lên

```
insert into SINHVIENGIOI
select MASV, HOTEN, DIEMTB
from SINHVIEN
where DIEMTB > 8
```

2.4 Cấu trúc CASE

Cấu trúc 1

```
CASE input_expression

WHEN when_expression THEN result_expression

[ ...n ]

[

ELSE else_result_expression
]

END
```

Ví dụ 7: Cho biết họ tên các nhân viên và năm về hưu

```
SELECT HONV, TENNV,

(CASE PHAI

WHEN 'Nam' THEN YEAR(NGSINH) + 60

WHEN 'Nu' THEN YEAR(NGSINH) + 55

END ) AS NAMVEHUU

FROM NHANVIEN
```

Ví dụ 8: Cho biết họ tên các nhân viên đã đến tuổi về hưu (nam 60 tuổi, nữ 55 tuổi)

```
SELECT HONV, TENNV
FROM NHANVIEN
WHERE YEAR(GETDATE()) - YEAR(NGSINH) >= ( CASE PHAI WHEN 'Nam' THEN 60 WHEN 'Nu' THEN 55 END )
```

Cấu trúc 2

```
CASE
WHEN Boolean_expression THEN result_expression
[ ...n ]
[
ELSE else_result_expression
]
END
```

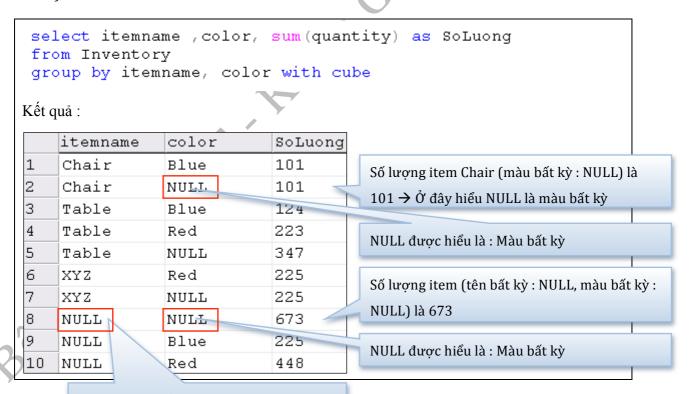
Ví dụ 9: Cho biết sinh viên và xếp loại học lực của sinh viên

2.5 GROUP BY ... WITH CUBE | ROLLUP với GROUPING

Trường hợp 1:

	itemid	itemname	color	quantity
1	1	Table	Blue	124
2	2	Table	Red	223
3	3	Chair	Blue	101
4	4	XYZ	Red	225

Ví dụ 10a. Cho biết số lượng item ứng với (mỗi tên, mỗi màu), (từng tên), (từng màu)



NULL được hiểu là : Tên bất kỳ

Trường hợp 2 (Sửa dữ liệu) → Vấn đề ???

Thay đổi dữ liệu ở bảng trên:

	itemid	itemname	color	quantity	
1	1	Table	Blue	124	
2	2	Table	Red	223	Có một item có
3	3	Chair	Blue	101	itemname giá trị
4	4	NULL	Nou		NULL

Kết quả thực hiện lại câu truy vấn:

	itemname	color	SoLuong	
1	NULL	Red	225	.1 >
2	NULL	***** *	225	NULL được hiểu là tên bất kỳ
3	Chair	Blue	101	hay là itemname = NULL →
4	Chair	NULL	101	
5	Table	Blue	124	?????
6	Table	Red	223	
7	Table	NULL	347	
8	NULL	NULL	673	
9	NULL	Blue	225	
10	NULL	Red	448	()

→ Không phân biệt giá trị NULL ở đây hiểu là gì : là tên bằng NULL hay ý nghĩa là tên bất kỳ → Giải pháp : Sử dụng Grouping để xác định

Sử dụng GROUPING:

Cú pháp : GROUPING (<tên thuộc tính>)

Grouping nhằm kiểm tra xem giá trị của một thuộc tính ở kết quả của câu truy vấn là do dữ liệu thực sự hay là do câu truy vấn tự phát sinh

Giải thích với ví dụ trên:

Nếu Grouping giá trị của thuộc tính tên (itemname) là $1 \rightarrow$ giá trị NULL xuất hiện do mệnh đề CUBE \rightarrow ý nghĩa là tên bất kỳ (do câu truy vấn tự phát sinh)

Nếu Grouping giá trị của thuộc tính tên(itemname) là $0 \rightarrow$ giá trị NULL có ý nghĩa là tên của item đó là NULL (do dữ liệu thực sự: itemname là NULL).

Ví du:

```
select itemname ,color, sum(quantity) as SoLuong,
grouping (itemname) 'grp'
from Inventory
group by itemname, color with cube
```

Kết quả là:

	itemname	color	SoLuong	grp
1	NULL	Pad	225	0
2	NULL	NULL	-	N
3	Chair	Blue	101	O N
4	Chair	NULL	101	0
5	Table	Blue	124	0
6	Table	Red	223	0
7	Table	NULL	347	0
8	NULL	NULL	673	1 N
9	NULL	Blue		G
10	NULL	Red	448	1

NULL được hiểu là itemname = NULL do Grouping (itemname)=0

→ Item NULL màu đỏ có số lượng là 225

NULL được hiểu là tên bất kỳ do Grouping (itemname) = 1

→ Item bất kỳ, màu đỏ có số lượng là 448

2.6 Sử dụng Case ... End với Grouping

Kết quả là:

	Itemname			Color			SoLuong
1	NULL			Red			225
2	NULL			Màu	bất	kỳ	225
3	Chair			Blue			101
4	Chair			Màu	bất	kỳ	101
5	Table			Blue			124
6	Table			Red			223
7	Table			Màu	bất	kỳ	347
8	Tên	bất	kỳ	Màu	bất	kỳ	673
9	Tên bất kỳ			Blue			225
10	Tên	bất	kỳ	Red			448

3 Bài tập tại lớp

<u>Yêu cầu</u>: Viết các câu truy vấn sau bằng ngôn ngữ SQL, sử dụng CSDL Quản lý đề tài.

- a. Cho biết danh sách tất cả giáo viên (magy, hoten) và họ tên giáo viên là quản lý chuyên môn của họ.
- b. Cho biết danh sách tất cả bộ môn (mabm, tenbm), tên trưởng bộ môn cùng số lượng giáo viên của mỗi bộ môn.
- c. Cho biết danh sách tất cả các giáo viên nam và thông tin các công việc mà họ đã tham gia.
- d. Cho biết danh sách tất cả các giáo viên và thông tin các công việc thuộc đề tài 001 mà họ tham gia.
- e. Cho biết thông tin các trưởng bộ môn (magv, hoten) sẽ về hưu vào năm 2014. Biết rằng độ tuổi về hưu của giáo viên nam là 60 còn giáo viên nữ là 55.
- f. Cho biết thông tin các trưởng khoa (magy) và năm họ sẽ về hưu.
- g. Tạo bảng DANHSACHTHIDUA (magv, sodtdat, danhhieu) gồm thông tin mã giáo viên, số đề tài họ tham gia đạt kết quả và danh hiệu thi đua:
 - o Insert dữ liệu cho bảng này (để trống cột danh hiệu)

- Dựa vào cột sldtdat (số lượng đề tài tham gia có kết quả là "đạt") để
 câp nhât dữ liêu cho côt danh hiệu theo quy đinh:
 - Sodtdat = 0 thì danh hiệu "chưa hoàn thành nhiệm vụ"
 - 1 <= Sodtdat <= 2 thì danh hiệu "hoàn thành nhiệm vụ"
 - 3 <= Sodtdat <= 5 thì danh hiệu "tiên tiến"
 - Sodtdat >= 6 thì danh hiệu "lao đông xuất sắc"
- h. Cho biết magy, họ tên và mức lương các giáo viên nữ của khoa "Công nghệ thông tin", mức lương trung bình, mức lương lớn nhất và nhỏ nhất của các giáo viên này.
- Cho biết makhoa, tenkhoa, số lượng gv từng khoa, số lượng gv trung bình,
 lớn nhất và nhỏ nhất của các khoa này.
- j. Cho biết danh sách các tên chủ đề, kinh phí cho chủ đề (là kinh phí cấp cho các đề tài thuộc chủ đề), tổng kinh phí, kinh phí lớn nhất và nhỏ nhất cho các chủ đề.
- k. Cho biết madt, tendt, kinh phí đề tài, mức kinh phí tổng và trung bình của các đề tài này theo từng giáo viên chủ nhiệm.
- l. Cho biết madt, tendt, kinh phí đề tài, mức kinh phí tổng và trung bình của các đề tài này theo từng cấp độ đề tài.
- m. Tổng hợp số lượng các đề tài theo (cấp độ, chủ đề), theo (cấp độ), theo (chủ đề).
- n. Tổng hợp mức lương tổng của các giáo viên theo (bộ môn, phái), theo (bộ môn).
- o. Tổng hợp số lượng các giáo viên của khoa CNTT theo (bộ môn, lương), theo (bộ môn), theo (lương).

Thời lượng: 02 giờ.

4 Bài tập về nhà

Yêu cầu:

1. Viết các câu truy vấn từ Q75 → Q82 của bài tập Quản lý Đề tài.

Bài tập **Quản lý ĐỀ TÀI**:

- Q75. Cho biết họ tên giáo viên và tên bộ môn họ làm trưởng bộ môn nếu có
- Q76. Cho danh sách tên bộ môn và họ tên trưởng bộ môn đó nếu có
- Q77. Cho danh sách tên giáo viên và các đề tài giáo viên đó chủ nhiệm nếu có
- Q78. Xóa các đề tài thuộc chủ đề "NCPT".
- Q79. Xuất ra thông tin của giáo viên (MAGV, HOTEN) và mức lương của giáo viên. Mức lương được xếp theo quy tắc: Lương của giáo viên < \$1800 : "THẤP"; Từ \$1800 đến \$2200: TRUNG BÌNH; Lương > \$2200: "CAO"
- Q80. Xuất ra thông tin giáo viên (MAGV, HOTEN) và xếp hạng dựa vào mức lương. Nếu giáo viên có lương cao nhất thì hạng là 1.
- Q81. Xuất ra thông tin thu nhập của giáo viên. Thu nhập của giáo viên được tính bằng LƯƠNG + PHỤ CẤP. Nếu giáo viên là trưởng bộ môn thì PHỤ CẤP là 300, và giáo viên là trưởng khoa thì PHỤ CẤP là 600.
- Q82. Xuất ra năm mà giáo viên dự kiến sẽ nghĩ hưu với quy định: Tuổi nghỉ hưu của Nam là 60, của Nữ là 55.
- 2. Viết các câu truy vấn sau bằng ngôn ngữ SQL, sử dụng CSDL Quản lý chuyến bay.
 - a. Tổng hợp số lượng các nhân viên theo (loại nhân viên, mức lương), theo (loại nhân viên) và theo (mức lương).
 - b. Tổng hợp số lần máy bay B747 và B727 đã cất cánh theo (ngày đi, mã cb) và theo (ngày đi).
 - c. Cho biết tên nhân viên, mức lương nhân viên, mức lương lớn nhất, nhỏ nhất và trung bình của các nhân viên này theo từng loại nhân viên.

- d. Cho biết tên các khách hàng có mã số khách hàng nằm trong khoảng 0009 đến 0400, số chuyến bay mà khách hàng đặt chỗ và tổng số chuyến bay mà tất cả các khách hàng này đặt.
- e. Cho biết tên nhân viên, tuổi nhân viên, độ tuổi trung bình của các nhân viên này theo mức lương nhân viên.
- f. Cho biết thông tin lịch bay của các chuyến bay (mã ch, ngày đi), số nhân viên trong phi hành đoàn (các nhân viên được phân công trên chuyến bay), số nhân viên ít nhất, nhiều nhất và trung bình của các phi hành đoàn.
- g. Tổng hợp số lần mà các nhân viên được phân công theo (mã cb, ngày đi), theo (ngày đi) và theo (mã cb).
- h. Tổng hợp số lượng chuyến bay theo (sbdi, sbden), theo (sbdi) và theo (sbden).
- i. Tổng hợp số lượng chuyến bay theo (giodi, gioden), theo (giodi) và theo (gioden).
- j. Cho biết tên nhân viên và mức phụ cấp cho nhân viên, biết rằng mức phụ cấp là 10% tiền lương nếu nhân viên là tiếp viên và 15% tiền lương nếu nhân viên là phi công.
- k. Cho danh sách tất cả nhân viên (manv, ten) và số lần họ được phân công trên các chuyến bay.
- l. Cho danh sách tên tất cả phi công và số loại máy bay họ có khả năng lái.
- m. Cho danh sách tất cả các loại máy bay và số lượng phi công có khả năng lái loại máy bay này.
- n. Cho danh sách tất cả các máy bay (mã loại, số hiệu) và số lần các máy bay này được đưa vào sử dụng cho các chuyến bay.
- o. Cho danh sách tất cả sân bay đi và số lượng chuyến bay đã cất cánh từ sân bay này vào ngày 11/1/2000.
- p. Cho danh sách mã và tên tất cả khách hàng cùng số lần mà khách hàng này đã đặt chỗ trên chuyến bay có mã 100.
- q. Cho danh sách tất cả ngày đi và số lượng máy bay của Airbus đã bay vào mỗi ngày này.

- r. Tạo bảng KhachHangThuongXuyen (makh, tenkh, solandatcho) chứa thông tin những khách hàng đã đặt chỗ từ 2 lần trở lên. Insert dữ liệu cho bảng này.
- s. Thêm cột loaikh vào bảng KhachHangThuongXuyen, cập nhật dữ liệu cho cột này theo quy định:
 - Nếu khách đặt chỗ từ 2 4 lần thì loại khách hàng là "thường xuyên"
 - Nếu khách đặt chỗ từ 5 7 lần thì loại khách hàng là "thân thiết"
 - Nếu khách đặt chỗ trên 7 lần thì loại khách hàng là "vip"

Thời lượng: 03 giờ.

ΗÉΤ