Data Warehousing and Data Mining

SQL Exercise (Non-Coursework)

1. Consider a database consisting of four tables **SUPPLIER**, **PART**, **PROJECT** and **SUPPLY** holding information on suppliers, parts, projects and the relationship between them: suppliers supply a certain quantity of parts to projects.

SUPPLIER

S#	SNAME	STATUS	CITY
S1	SMITH	20	LONDON
S2	JONES	10	PARIS
S3	BLAKE	30	PARIS
S4	CLARK	20	LONDON
S5	ADAMS	30	ATHENS

PART

P#	PNAME	COLOUR	WEIGHT	CITY
P1	NUT	RED	12	LONDON
P2	BOLT	GREEN	17	PARIS
P3	SCREW	BLUE	17	ROME
P4	SCREW	RED	14	LONDON
P5	CAM	BLUE	12	PARIS
P6	COG	RED	19	LONDON

PROJECT

J#	JNAME	CITY
J1	SORTER	PARIS
J2	DISPLAY	ROME
J3	OCR	ATHENS
J4	CONSOLE	ATHENS
J5	RAID	LONDON
J6	EDS	OSLO
J7	TAPE	LONDON

SUPPLY

S#	P#	J#	QUANTITY
S1	P1	J1	200
S1	P1	J4	700
S2	P3	J1	400
S2	P3	J2	200
S2	P3	J3	200
S2	P3	J4	500
S2	P3	J5	600
S2	P3	J6	400
S2	P3	J7	800
S2	P5	J2	100
S3	P3	J1	200
S3	P4	J2	500
S4	P6	J3	300
S4	P6	J7	300
S5	P2	J2	200
S5	P2	J4	100
S5	P5	J5	500
S5	P5	J7	100
S5	P6	J2	200
S5	P1	J4	100
S5	P3	J4	200
S5	P4	J4	800
S5	P5	J4	400
S5	P6	J4	500

A user of this database formulates SQL expressions for queries (i) to (iv) as follows. Unfortunately each has one or more syntax errors or fails to produce the correct result.

(i) Find the name and city of each project supplied with a red part.

```
SELECT JNAME, CITY
FROM PROJECT, PART, SUPPLY
WHERE COLOUR = 'RED'
```

(ii) Find the name and city of each project not supplied with a red part.

```
SELECT JNAME, CITY
FROM PROJECT, PART, SUPPLY
WHERE COLOUR <> 'RED'
```

(iii) For each red part, find the total number of suppliers supplying the part.

```
SELECT P#, COUNT(*)
FROM PART P, SUPPLY SPJ
WHERE P.P# = SPJ.P#
GROUP BY P#
HAVING COLOUR = 'RED'
```

(iv) Find the name and city of each supplier supplying every red part.

```
SELECT SNAME, CITY
FROM SUPPLIER S, SUPPLY SPJ
WHERE S.S# = SPJ.S#
GROUP BY S#
HAVING COUNT(*) = (SELECT *
                   FROM PART
                   WHERE COLOUR = 'RED')
```

For each query, work out what the errors are in the SQL statements and modify them to produce the correct results. The result rows which should be retrieved are shown below.

(i)

JNAME	CITY
SORTER	PARIS
DISPLAY	ROME
OCR	ATHENS
CONSOLE	ATHENS
TAPE	LONDON

(ii)

JNAME	CITY
RAID	LONDON
EDS	OSLO

(iii)

P#	TOTAL NUMBER
P1	2
P4	2
P6	2

(iv)

SNAME	CITY
ADAMS	ATHENS

2. Consider a database consisting of two tables UKCONSTS and UKRESULTS holding UK election data for a particular general election.

   Each UK constituency gives rise to a row in UKCONSTS with the following information:
   - an identifying number UKNUM
   - the constituency name UKAREA
   - the number of registered voters in the constituency UKELECTORS

   The result of each party in each constituency in the 1983 general election gives rise to a row in UKRESULTS with the following information:
   - the constituency number UKNUM
   - the party PARTY
   - the number of votes for that party UKVOTES

   For this election database, formulate SQL queries for the following. The Oracle function `ROUND(n, integer)` has been used to round value `n` to `integer` decimal places.

   (i) For each constituency in which more than 5 parties had candidates, find the average vote received by parties in that constituency. Order the rows by constituency.

UKAREA	AVERAGE VOTE
antrim east	6368
antrim south	6471.5
ashford	7989.33
bath	7978.33
battesea	6258.71
belfast north	6059.86
etc.	

   (ii) Find the constituency name, number of voters voting, and turnout (i.e. percentage of registered voters who voted) in constituencies where less than 65% of registered voters voted. Order the rows by ascending turnout.

UKAREA	TOTAL VOTE	TURNOUT
bethnal green and stepney	30838	54.39
battesea	43811	60.88
barking	34220	61.32
belfast east	37328	63.05
belfast south	37358	64.48
belfast north	42419	64.54
etc.		

(iii) Find the constituency name and turnout in constituencies which have at least 5000 more registered voters than the average number of registered voters in a constituency. Order the rows by descending turnout.

UKAREA	TURNOUT
berkshire east	89.92
bedfordshire mid	85.26
aldershot	78.90
bedfordshire north	78.82
beverley	78.79
arundel	77.50
etc.	

(iv) Find the average turnout in constituencies.

AVERAGE TURNOUT
74.36

(v) For each party, find the total number of seats won, i.e the number of constituencies in which they were the winning party. Order the rows by descending total number of seats.

PARTY	TOTAL SEATS
con	29
lab	13
oup	4
dup	2
lib	1
sf	1