
SWINBURNE UNIVERSITY OF TECHNOLOGY

FACULTY OF BUSINESS AND ENTERPRISE

INF30010/INF80037 DATABASE ADMINISTRATION

QUERY OPTIMISATION ASSIGNMENT

Date Due: 5.30pm on 21 September 2015

Assessment: 20% of total assessment

The assignment is to be done in teams of 1 to 2. Each team (including teams of 1) must be registered on ESP. The registration must be complete by Monday 14 September 2015.

Requirements:

You are to examine index strategies and the use of Inter File Clustering to make the following query perform efficiently from the point of view of total query cost and also the cost of access to the first row using a tablespace with a block size of 4096 and another tablespace with a block size of 2048:

```
select  vintage, wine_no, wname, pctalc, grade, price, wine.vid,  
        vname, wine.cid, cname  
  
from    vineyard, class, wine  
  
where   wine.vid = vineyard.vid  
  
and     wine.cid = class.cid  
  
and     wine.cid = 'SHIRAZ' and grade = 'A' ;
```

There are two versions of the database tables in the dba database on Oracle. The WINE, VINEYARD, PRODUCER, etc. tables are created in the USERS4K tablespace with a block size of 4096. The WINE2, VINEYARD2, PRODUCER2, etc. tables are for the USERS2K tablespace with a block size of 2048. **You have to run the scripts provided on BB for the Pass and Credit level work.**

=====

PASS GRADE

=====

To achieve a Pass in this assignment, you should examine a number of cases where you create an index (on one or more columns) or a combination of indexes to see **whether the optimiser chooses to use your indexes. The only hint you can use to generate test cases is first rows(1). At this stage there should be no Index Hints.**

1. Examine the query to determine the columns where it might be useful to create an index to improve performance.
2. Work out all the possible combinations of these indexes and write each of these index cases in the first column of the results table.
3. Use the version of QPLAN to look at the query plan for both all_rows (the default) and QPLAN_FR for first_rows(1). Give every **different** plan you see a name such as PLANa, PLANb. Fill in the results table with plan names and costs.
4. Copy each different plan you find with its name (as indicated in the table) in the space after the table.
5. Repeat steps 3 and 4 using the tables WINE2, VINEYARD2, CLASS2 etc. that have been created in tablespace USERS2K. Note that scripts CTSTATS2, QSTATS2, QPLAN2 and QPLAN2_FR have been provided to assist you with this process.
6. Write a conclusion section where you discuss in which cases Oracle used an index you created and when it did not, giving possible reasons and making recommendations in which cases it is advantageous to create an index.

RESULTS TABLE FOR BLOCK SIZE 4096

<u>File Name For Results</u>	<u>Index Case</u>	<u>All rows plan</u>	<u>Cost</u>	<u>First rows(1) plan</u>	<u>Cost</u>
PLANA4K/ PLANA4KFR	PK indexes only	PLANA4K	28	PLANA4K	5
	Wine(CID)				
	Wine(grade)				
...					

RESULTS TABLE FOR BLOCK SIZE 2048

<u>File Name For Results</u>	<u>Index Case</u>	<u>All rows plan</u>	<u>Cost</u>	<u>First rows(1) plan</u>	<u>Cost</u>
PLANA2K/ PLANA2KFR	PK indexes only	PLANA2K	40	PLAN2KA	5
	Wine(CID)				
	Wine(grade)				

CREDIT GRADE

Additional Requirements:

To achieve a Pass Grade, you examined a number of index cases to see if the optimiser would use the indexes in its plan for a given query. There were a number of cases where there was a single index on a single column and the optimiser chose not to use the index. For a credit grade, you extend this work by using the index hint to force the optimiser to use the index and examine the resulting plan. This should be repeated for each tablespace block size.

Document the resulting query plans in the same way as for a Pass grade: The overview table states the names of the plans used and after the table you list the distinct plans (do not duplicate plans).

Having done this you have to evaluate each of the different plans you have seen in each of the two stages. For each of these plans (both all_rows and first_rows) and for each of the tablespaces (USERS4K and USERS2K), you are to do your own cost calculation and show your results in the following format:

	Cost of step	Explanation
QUERY_PLAN		

0_ SELECT STATEMENT Cost = 28		
1__ HASH JOIN Cost = 28		
2___ TABLE ACCESS FULL WINE Cost = 21		
3___ NESTED LOOPS Cost = 6		
4_____ TABLE ACCESS BY INDEX ROWID CLASS Cost = 1		
5_____ INDEX UNIQUE SCAN SYS_C0084315 Cost = 0		
6_____ TABLE ACCESS FULL VINEYARD Cost = 5		

For the cost of step which requires a calculation you should show and explain your workings.

Note that these calculations are only estimates. We are more interested in your understanding of the plan. This step should be completed for both TABLESPACES i.e. block size 2048 and block size 4096.

DISTINCTION / HIGH DISTINCTION GRADE

Additional Requirements:

Explore the extended query

```
select  vintage, wine_no, wname, pctalc, grade, price, wine.vid,
vname, wine.cid, cname

from    vineyard, class, wine

where   wine.vid = vineyard.vid

and     wine.cid = class.cid

and     wine.cid = 'SHIRAZ' and grade = 'A' AND pctalc = 12.6 ;
```

with and without an interfile cluster.

The scripts are provided on Blackboard in the same folder as this task sheet.

Make a decision which tables should be included in the cluster. For a high distinction, explore plausible options and compare the resulting query plans (with and without a cluster).

Also compare the costings of the following queries before and after creating the cluster:

```
select  vintage, wine_no, wname, pctalc , grade

from    wine

where   wine.cid = 'SHIRAZ' and grade = 'A' ;
```

```
select * from class

where colour='RED' ;
```

```
select * from vineyard where pid='PEN' ;
```

Document the plans as before, with overview table and listing of distinct plans.

Write a conclusion section where you discuss the query plans (why the queries were executed in the way they were) and give a recommendation whether the cluster(s) should be used in the WINE database.

REPORT REQUIREMENTS:

The following items need to be included as a part of your final submission:

1. A document title with your name and student number.
2. An Introduction where you state which grade you are aiming for and indicate the sections which you have completed for each grade.
3. An appropriate heading for each section (e.g. Query Plans for Block Size 4K without Index).
4. The sections needed for each grade as outlined above, with the appropriate conclusions.
5. A conclusion for the whole report where you discuss what you have learned about the use **of indexes** (for higher grades **also clusters**) to make queries more efficient (when they are useful and when not).

Convert the completed document to .pdf format and submit through ESP.

<https://esp.ict.swin.edu.au/>