CSC2001F - ASSIGNMENT 4 - 2019

Three CSV files are supplied that store (partly fabricated) data about people and their spending. All amounts are thus stored as negatives, but you must use only their absolute value in your queries and their outputs. ID is a unique identifier for a person.

Create a MySQL database to store this data in relations with the exact structure and content of these CSV files – call these relations Spend, CodeCat and Account. Then write SQL queries to answer the questions below and run them against your database. Do this **individually, not in groups**. Your queries must be correct for any instance of this database schema, and not just for the given sample data. If MySQL is unable to handle any of the queries, state this clearly in your answer, with a brief reason why. If you cannot find the precise information required, give information as close as possible to what is required, explaining clearly (in words) what it is you are providing. If you believe a request is ambiguous, explain why, and state the interpretation you selected. **Please use the Vula Forum for questions** (emails will NOT be answered) so everyone has the same information.

Submit all your answers to this assignment in 1 PDF file.

Section A   [5 marks]

a) For each of the 3 relations state what attribute(s) you would use as its primary key.
b) List all foreign keys in your 3 relations, in the form *relationName.attributeName*

Section B   [32 marks]

Give **question number, question text, SQL code and output** for each request in turn :

1. Show any one complete tuple (just 1, no more) from each of your 3 relations.
2. Output the number of tuples in each of your 3 relations.
3. List first name and surname of everyone with surname "Naidoo", in alphabetical order of first name.
4. For each code 5511 transaction, give the transactionDate along with the VAT on that transation, where the VAT was 15% of the amount spent (e.g. if the amount was -115 then the VAT was 15).
5. Show the firstname, surname, age, marital-status and income category of everyone who has any missing data in any of these fields.
6. Give the names of merchants who have had transactions of both codes 5111 and 2741.
7. Find the transactionDate and merchant associated with the highest amount in the database, along with that amount.
8. How many people (ids) have any transaction amount that is higher than every amount ever spent by the person with id 12312870316 ?
9. How many categories are there? (i.e. how many different category values?)
10. Give the average transaction amount (a single value in the answer).
11. Find the code range for each category (i.e. each result line will have category

name, lowest code and highest code for that category).
12. Find the largest transaction amount for each code that is associated with more than 50 transactions.
13. Which code(s) have the least number of transactions i.e. which code(s) have the fewest transactions?
14. Show all information on all large transactions including the category involved. A large transaction is one with an amount that exceeds ten times the average transaction amount (e.g. if the average transaction amount is R100, a large transaction has an amount above R1000).
15. Find the total amount spent on Pets for each merchant where there have been more than 10 transactions in the Pets category.
16. For each category, show how many such transactions there are in the database along with the total Rand amount of those transactions.

Section C   [5 marks]

Devise a useful query of your own involving the most interesting SQL usage you can think of. Explain clearly what it is meant to find out from the database, then show the SQL and its result, and then show/explain clearly how you know it is correct. Marks are awarded based on how challenging the query is and how clearly its correctness is shown.

Section D   [8 marks]

 I.   Delete all transactions from merchant 3@1 (however it may be spelt).
 II.   Insert a new transaction for id 12393560590 that took place at midnight on New Years Eve 2018 for an amount of R448. The code is 5251 but the merchant and description are not known.
 III.   Code 5211 purchases should have been coded 5221, so change 5211 to 5221.
 IV.   For everyone who is not M(married), add another row for them, which has their maritalStatus as U (unmarried) and has the values in all the other columns of their new extra row the same as in their original row.