INFO20003 Semester 2, 2021

Assignment 2: SQL
**Due**: 17/09/2021 at 6pm
Submission: Via LMS https://canvas.lms.unimelb.edu.au/

Weighting: 10% of your total assessment

Contact Centre case study

The case study focuses on a small part of a contact centre operated by a large financial institution. The contact centre employs several staff members organised in a hierarchical team structure. The case study focuses on four teams which are named after Australian national parks and are organised as follows:
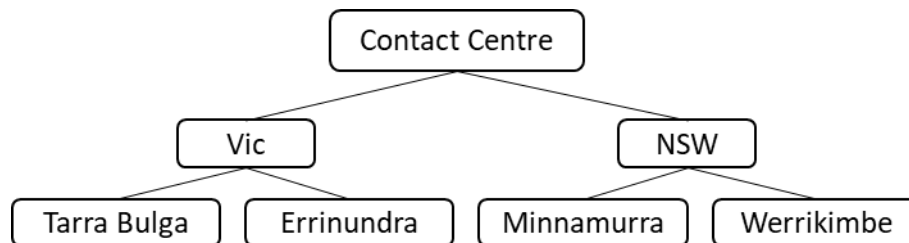
Each team has a team leader and a small number of agents (between 1 and 4). The team structure of contact centres is very dynamic in real life – and this case is no exception. Staff members may move from one team to another, may be scheduled to work for more than one team at the same time and may assume different roles within the same or different teams.

The *profile* table keeps track of staff allocations over time using a unique 4-digit identifier for each allocation of a member of staff to a team in a particular role (e.g. Team Leader or Agent). The *valid_from* and *valid_until* attributes track the duration of the allocation.

Throughout the day, Contact Centre agents take incoming calls from customers. For every call, the system generates an entry in the c*all_record* table, linking it to the agent who handled the call. A call may be transferred between agents or agents to team leaders if need be. In that case, the call comprises of more than one call legs. The provided call record data covers a period of three months from 1st April 2021 to 30th June 2021.

Important note: you may use the time of the survey (survey.response_time) when determining 'which month was a call+survey response made in', as opposed to the time the call was first started (call_record.call_time).

Contact Centre management requires that at the end of every call, the agent asks the caller to stay on the line and complete a post-call survey using their phone keypad. The survey responses are used to analyse the performance of contact centre individuals and teams, as well as the overall business. The survey consists of three questions, as follows:

#	Question Metric	Question Wording	Answer
1	First Call Resolution	Was your enquiry resolved today?	1-3 scale, where 1=yes, 2=no, 3=too early to tell
2	Agent Quality	How happy were you with the person you spoke to?	0-10 scale, where 0=very unhappy and 10=extremely happy
3	Promoter Score	How likely are you to recommend us to your friends and family?	0-10 scale, where 0=highly unlikely and 10=highly likely

A caller may opt out of taking a survey altogether. Some callers may agree to the survey but quickly hang up without answering any questions. Others may answer only some of the questions before hanging up. And some will give answers to all questions. Customers cannot 'skip' questions, i.e. they can't have provided an answer to Q2 without having provided an answer to Q1. Customer responses are recorded in the s*urvey_response* table.
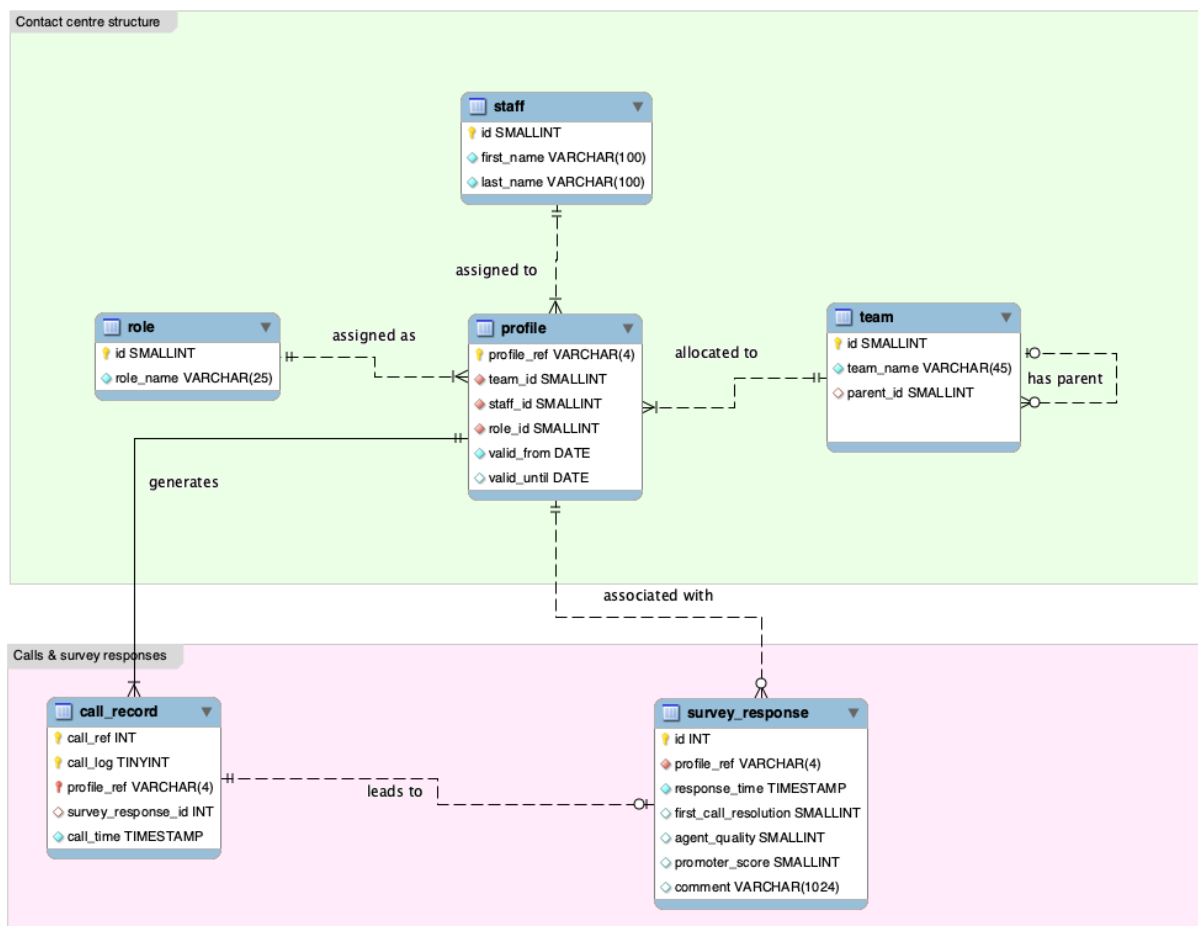
For reference, the data model is given below.

Figure 1: The ER Model for Contact Centre Database

Some definitions

- "Participation" is a contact centre metric that measures what proportion (percentage) of inbound calls are followed up by an attempted survey. This number can go anywhere between 0% and 100%. Contact centre agents and management would like this figure to be as high as possible.
- "Enhanced participation" is like participation, but considers not just offered surveys, but offered surveys where at least the first question was answered by the customer (i.e., in this case they didn't "say yes to the survey, but then hang up before answering the first question").
- "Net Promoter Score" or NPS is a measure of customer loyalty and is calculated from the collected promoter scores. The way how NPS is calculated can be found here https://www.netpromoter.com/know/. An online calculator is available on http://www.npscalculator.com/en#
- "Cherry picking" refers to the practice of agents selectively offering surveys only to callers who are likely to score the agent highly – with a 9 or 10 – on the Agent Quality question.

Assignment 2 Setup

A dataset is provided against which you can test your solutions to the assignment. To set up the dataset, download the file *A2_sqlconfig_INFO20003.sql* from the Assignment on Canvas and run it in MySQL Workbench. This script creates the database tables and populates them with data.

The script is designed to run against your account on the Engineering IT server (info20003db.eng.unimelb.edu.au). If you want to install the schema on your own MySQL Server installation, uncomment the three lines at the beginning of the script that create the schema on your local server.

Note: Do NOT disable full_group_by mode when completing this assignment*. This mode is the default, and is turned on in all default installs of MySQL workbench. You can check whether it is turned on using the command "SELECT @@sql_mode;". It should return a string containing "full_group_by". When testing, our test server WILL have this mode turned on, and if your query fails due to this, you will lose marks. You can run the command:*

> *SET sql_mode=(SELECT CONCAT(@@sql_mode,',ONLY_FULL_GROUP_BY));*

… to ensure that this mode is configured properly, you must run this every time you start mysql.

The SQL Tasks

In this section are listed 10 questions for you to answer. Write one (single) SQL statement per question. Subqueries and nesting are allowed within a single SQL statement. However, you may be penalized for writing *overly* complicated SQL statements. DO NOT USE VIEWS (or 'WITH' statements/common table expressions) to answer questions.

1. Display the first and last names of all the team leads. Do not repeat names if a staff member has been a team lead more than once. Return as (first name, last name).

   (1 mark)

2. Show all staff who have ever worked for a Victorian team. Show their first and last name and team. Display the list sorted alphabetically by last name. Return as (last name, first name, team name).

   Note that some staff may have worked for multiple teams. In that case, their name should appear in multiple records, each time belonging to a different team.

   (1 mark)

3. Show staff who worked in Errinundra at any time between 13th April and 13th May. Return as (first name, last name). Do not show duplicate records.

   (1 mark)

4. Which agents (not team leaders) were working only in Werrikimbe and did not work as *agents* in any other team on 23rd of May 2021? Return as (first name, last name).

   (2 marks)

5. Calculate the monthly average agent quality (AQ) score for each team. Display the team's name, month and average AQ sorted by average AQ so that the highest scoring teams are listed first. Return as (team name, month, averageAQ).

   (2 marks)

6. The business wants to know how their Net Promoter Score (NPS) is tracking. Calculate the NPS of the organisation for the months of April, May, and June 2021. Surveys without a valid promoter score (ie customers skipped the last question) should be omitted from the calculation. Return as (month, NPS) sorted by month ascending.

   (2 marks)

7. What was Tatjana Pryor's enhanced participation in June? Return as (Tatjana's enhanced participation).

   (2 marks)

8. What agent has the lowest number of First Call Resolution (FCR) is "no" survey responses in the period 01 June - 17 June? If there are multiple agents with equal values, return them all. Return as (first name, last name, FCR count).

   (3 marks)

9. Show the average agent quality (AQ) score for interactions where the agent has needed to contact a team leader during the interaction (i.e., a leg of the call involved a team leader), compared to the overall average AQ. Return as (average AQ when a team leader was involved, overall average AQ).

(3 marks)

10. *"Wait, you've never met Tatjana?"* Find the names of staff members that have worked in ALL the teams EXCEPT for the teams that 'Tatjana Pryor' (staff id = 2) has worked in. Only consider teams which 'have staff' (i.e. has_staff = 1). Return as (staff first name, staff last name).
Hint: To help you get started, this is an example of a 'relational divide'.

(3 marks)

SQL Response Formatting Requirements

To help us mark your assignment queries as quickly/accurately as possible, please ensure that:

- Your query returns the projected attributes in the same order as given in the question, and does **not** include additional columns. E.g., if the question asks 'return as (userId, name)', please write `"SELECT userId, name …"` instead of `"SELECT name, userId…"` (*you can name the columns using '`AS`' however you'd like, only the **order** matters)*

- Please do NOT use "databaseName.tableName" format. E.g., please write `"SELECT userId FROM users…"` instead of `"SELECT userId FROM coltonc.users …"`

- Ensure that you are using single quotes(') for strings (e.g. `…WHERE name = 'bob'…`) and double quotes (") or backticks (` ) only for table names (e.g. `SELECT name FROM `some table name with spaces`…`). Do **NOT** use doublequotes for strings `"…WHERE name = "bob"…"`.

- Ensure that you match the capitalisation of table/attribute names: some OS's are case insensitive but others are case sensitive. E.g. for an attribute "name" in table "user", please write `"SELECT name, FROM user …"` instead of `"SELECT Name FROM User…"`

Submission Instructions

Your submission will be in the form of an SQL script. There is a template file on the LMS, into which you will paste your solutions and fill in your student details (more information below).

This .sql file should be submitted on Canvas by **6pm** on the due date of **Friday 17 of September**. Name your submission as 987654.sql, where 987654 corresponds to YOUR student id.

Filling in the template file:
The template file on the LMS has spaces for you to fill in your student details and your answers to the questions. There is also an example prefilled script also available on the LMS. Below are screenshots from those two documents explaining the steps you need to take to submit your solutions:

Step	Example

1. At the top of the template, you'll need to replace "XXXXXXXX" with your student number and name	**Template**
	
	`-- Your Name: XXXXXXXX`
	`-- Your Student Number: XXXXXX`
	
	Example Filled in
	
	`-- Your Name: Colton Carner`
	`-- Your Student Number: 693281`

2. For each question 1-10, place your SQL solution in between the "BEGIN QX" and "END QX" markers. **Ensure each query is terminated with a semicolon ";"**	**Template**
	
	`-- _____`
	`-- BEGIN Q1`
	
	`-- END Q1`
	`-- _____`
	
	Example Filled in
	
	`-- _____`
	`-- BEGIN Q1`
	
	`-- **This is an example of a comment which will not be executed`
	`-- **(comments are not NEEDED, but if your query is complex you can leave some)`
	
	`-- **Below is an example of how you would enter your answer:`
	`SELECT *`
	`FROM delivery`
	`NATURAL JOIN deliveryitem`
	`WHERE supplierid = 101;`
	
	`-- **It's OK to add more space in between the 'BEGIN QX' and 'END QX' markers`
	`-- **for each question, just don't DELETE the markers!`
	
	`-- **Make sure you fill out your name + student num at the top of the document`
	
	`-- **After reading / understanding these comments in the Q1 section`
	`-- **(comments starting with '**'), feel free to delete them`
	`-- **(don't delete lines without **)`
	
	`-- END Q1`
	`-- _____`

3. Test that your script is valid SQL by running it from MySQL Workbench. Run the entire script by either a) using 'run sql script' from the file menu (can sometimes be a bit buggy) or b) placing your cursor at the start of the file (without selecting anything), pressing the lightning bolt to run the entire file, and checking that all queries ran successfully (all green ticks).	Running the example filled in the template provided, using the schema we've used in labs (dep-store) results in 'success'. (You would run yours on the schema for the A2 assignment)

Late submission

Unless you have an approved extension (see below), you will be penalised -10% of the total number of marks in the assignment per day that your submission is late. For instance, if you received a 78% raw score, but submitted 2 days late, you'd receive a 58% score for the assignment.

Requesting a Submission Deadline Extension

If you need an extension due to a valid (medical) reason, you will need to provide evidence to support your request by 5 pm on 16th of September. Medical certificates need to be at least two days in length.

To request an extension:

- Email Farah Zaib Khan (farah.khan@unimelb.edu.au) from your university email address, supplying your student ID, the extension request and supporting evidence. Please add INFO20003 in the subject title.
- If your submission deadline extension is granted you will receive an email reply granting the new submission date. Do not lose this email!

Reminder: INFO20003 Hurdle Requirements

To pass INFO20003, you must pass two hurdles:

- **Hurdle 1:** Obtain at least 50% (15/30) for the three assignments (each worth 10%)
- **Hurdle 2:** Obtain at least 50% (35/70) for the combination of the quizzes and final exam

Therefore, it is our recommendation that you attempt every assignment and question in the exam.

GOOD LUCK!