# MMAI5100 Group Project – Fall 2021

Project Objectives

Complete the conceptual, logical and physical design as well as the physical implementation of a relational database for a given business scenario.

Business Scenario

Realtek Inc. is a residential real estate property firm that represents clients for sale or rental of their residential properties. While its headquarters is located in Toronto (ON), Realtek Inc. has many offices across Canada.

You have your own consulting practice and lead a small team of consultants. You have recently been approached by Bobby Lau, Realtek Inc.'s COO who is a friend of yours from Schulich School of Business to design an operational database for Realtek's business needs. In your initial conversation with Bobby you learned that Realtek Inc. has gone through an extensive period of growth during the real estate boom in Canada over the last decade but failed to transform their IT operations to keep up with the expansion. It looks like the company still operates as it did originally (when it only had a single office in Toronto). When you asked him about existing databases and data pipelines, to your shock Bobby told you that Realtek still keeps track of its many offices' operations in a series of spreadsheets that get sent periodically via emails to the Headquarters.

In a subsequent meeting with the Realtek's upper management, your team was able to gather the following details about Realtek operations:

Each office is either owned/operated by the Realtek Inc. or is leased[1]/operated by an affiliated realtor (or a number of realtors). Each office has an Office ID and Office Address which indicates Street, City, Province and Postal Code for the office. Leased offices will have Lease Start and Lease End dates that are tracked by the headquarters.

Realtek Inc. has many partners which are themselves realtors and each corporately owned/operated office will have a Responsible Partner who would be responsible for that office. These corporately managed offices will be staffed by one or more full time employees (FTEs). Each corporate employee will have an Employee ID and an Employee Name (First Name, Middle Initial and Last Name). Each corporate employee can only work in one office. Corporately owned/operated offices will be managed by one of these employees (an employee can only manage the office he/she is currently working in).

Leased (affiliate operated) offices will have one or more affiliates (lease holders). Affiliates are tracked by their Name (First Name, Middle Initial and Last Name) and date of birth (DOB). There is no restriction in how many offices affiliates can lease so it is possible for an affiliate (or a bunch of affiliates) to operate number of offices. For leased offices, the leases are usually negotiated every 3 to 5 years so it is important to keep track of lease start/end dates.

Both corporately owned/operated and affiliate leased/operated offices will list on MLS (multiple listing service) residential properties for sale and rental. Each listed property will have a unique Property ID

1 Leased offices operate like franchises and pay royalty commissions to Realtek on their successful sales/rental deals.

and Property Address (Street, City, Province and Postal Code). Rental properties will have monthly rent, rental agreement period and commission amount while properties for sale will have a sales price, closing duration and commission percentage. Leased/operated offices pay a royalty commission of 10 percent to Realtek Inc. out of of the sales or rental commissions they make on successful deals. A residential property can be listed for both sale and rental at the same time. Each residential property can only be listed by one office whereas each office may list any number of properties (or none).

Each property will have one or more owners. Owners will have an Owner ID, Owner Name (First Name, Middle Initial and Last Name) and Owner Address (Street, City, Province and Postal Code). Owners may have one or more properties listed. On occasions when there are multiple owners, owners' ownership percentage will be tracked (Naturally, a single owner would have 100% ownership).

Presently, all data communication between the Headquarters and owned or leased offices rely on spreadsheets that are prepared by each office in regular intervals and transmitted via email. This kind of communication is not ideal as it creates extra work for each department to manually compile aggregate reports. It is not unusual for email transmissions be misplaced or accidentally deleted. There were even occasions when, due to mislabelling, spreadsheets were double counted in reports which led to accounting errors.

Your preliminary interviews with Realtek operational staff exposed that different departments had different reporting needs. For example:

HR needs regular (weekly) reports showing which staff worked in what office, who the operating manager is along with the responsible partner for that office. HR indicated it would be great if these reports can be organized or ordered in some fashion as current spreadsheet reports are not ordered in any way which causes them extra work to merge multiple spreadsheets to see details like which office has the largest or smallest workforce.

Legal needs a monthly list of all leased offices, who the lessor(s) are and in particular a list of leases that will be expired within the next 3 months so that they can work on renewal agreements if applicable.

Accounting needs weekly reports showing all kinds of sales and rental data for each office location including but not limited to total sales/commissions, total rentals/commissions per office as well as top and bottom 3 offices in sales and rentals commissions. A separate report is weekly prepared for royalty commissions received from leased offices. While accounting has a separate payroll application, they also would like to be able to keep track of employee income (salaries or commissions) details in the database.

Management needs a few reports to show them which locations are the most efficient (defined as commissions generated per employee) or the most profitable (collects the highest real estate commission percentages). For example, one of those reports could show average commissions earned per employee per month for both corporately owned and leased (affiliate operated) locations so that they can compare staff performance in each type of office in similar geographies.

Based on the information gathered your team decided that a centralized relational database would be the best and simplest technology option to serve Realtek's operational needs for the foreseeable future.

Due to its stellar reputation for performance, security and scalability and not to be burdened by exuberant licensing fees, the Team also decided that PostgreSQL RDBMS would be the preferred implementation platform.

Deliverables

-Prepare an E-R or an EER diagram to show the conceptual design for Realtek Inc.'s data needs.
-Convert your conceptual diagram into a logical data model for a relational data management technology implementation (transform your E-R or EER diagram into relations).
-Prepare a physical data model showing the data types for PostgreSQL DBMS implementation.
-Implement your design in PostgreSQL using Structured Query Language on our database server (available at mmai5100postgres.canadacentral.cloudapp.azure.com). A single SQL implementation is sufficient as the final deliverable (i.e. each group member does not have to duplicate the SQL code). Make sure you have sufficient sample data in the database (a few records should be sufficient in each table) so that the requirements can be tested and the queries will run.
-Please provide me with a single or a series of SQL scripts so that I can create and populate your database from scratch in my own database environment.
-As the basis of suggested reports, prepare a series of queries that will serve Realtek's day-to-day HR, legal, accounting as well as operational managerial needs. In your report, explain in detail what each query does and have the queries ready in a clearly commented script (or a series of scripts).
-Based on your knowledge of databases and ML/AI, propose future extensions to the current database work that may help Realtek maintain its competitive advantage in the market (and provide future business to your consulting company).

For your diagrams, you can use an online or desktop tool to create them or draw by hand, take a photo and insert the photo into your answer. If you are using a tool please pick a tool that has similar notation to the one we used in class. In any case, please make sure your relationship cardinalities are clearly visible.

For details that are not clearly stated in the case you are allowed to make assumptions. Nevertheless, if and when you make any assumptions, please state them clearly.

A GitLab instance (https://mmai5100gitlab.canadacentral.cloudapp.azure.com/) is made available for your project. You can use GitLab as your project management tool or at least as an overall repository for all your work. GitLab can create high-level ER diagrams but many online tools may be easier to use.

Finally, please note that there are numerous ways this database can be designed and constructed. If you are torn between multiple design options, opt for the simplest design and justify your selection.