# Data Assessment

## Objective

The objective of this individual assessment is to build OLTP and OLAP systems based on the learning outcome of Module #2

## Overview - Scenario

### ‘MyHomes’ Real Estate Company

MyHomes is a real estate company operating globally, their main area of business lies in buying and selling properties.

The properties that the company deals in are of five types: houses, flats/apartments, bungalows, land, and commercial property. These properties may be subject to health and safety regulations, which govern how these properties must be build and maintained.

Each property that MyHomes deals in is the main responsibility of one of the company’s area sales representatives. Each representative, however, may be responsible for a number of different properties. A number of the area sales representatives act as supervisors. Supervisors are directly responsible for managing teams of fellow area sales representatives. An area sales representative may also be responsible for dealing with a number of MyHomes’s customers. Each of the company’s customers will only deal with MyHomes through one particular sales representative. Each sales representative also covers one or more geographical sales areas, for eliciting new business from prospective customers. A number of different sales representatives may cover the same geographical area, but there will always be a single sales representative who acts as the chief salesperson for that area, and who is directly responsible to management for sales performance in that area, in terms of hitting sales targets.

You are asked to design and implement information systems (OLTP and OLAP) for MyHomes company.

## Tasks of the Assessment

**Building Online Transaction Processing System - Create a new database**

**Part 1 (20 points)**

*(A) Using a database design (relational model)/Normalization approach of your choice, produce a Physical design for the database to support the information system, which is needed at the* ***MyHomes.***

(10 points)

*(B) Based on your Physical design from Part 1 (A) and the information available in the scenario, produce an SQL script that create appropriate tables*

(10 points)

**Part 2 (30 points)**

This part is based on your answer / solution to Part 1, i.e., design and implementation of the database for the 'MyHomes' scenario.

*(A) Populate the database tables with some data*

(20 points)

*(B) Answer the following queries (retrievals) using SQL*

1) *Display names of representatives, details of the properties they represent, and names of their supervisors.*

2) *Display details of customers together with details of their areas and names of the managers of their areas.*

(10 points)

**Part 3 (50 points)**

**Online Analytical Processing System – Create a separate database**

*A) Designing Dimensional Modelling*

*Using Kimball's methodology, design a dimensional model(s) MyHomes's scenario that will allow the company to investigate as to why some cities in various countries have good sales whilst others have not. You may make any reasonable assumptions to support your design (state any assumptions you make).*

(10 points)

B) *Based on your dimensional model from Part 3 (A) produce an SQL script that create appropriate tables*

(5 Points)

*C) Move appropriate data from OLTP to OLAP system, consider applying ETL operations (Use Python Programming and relevant modules)*

(25 Points)

*D) Use Matplotlib/Data Visualization Tool to analyse and discover various associations within the given data set and demonstrate various types of graphical representations. Justify your analysis approach and methods/tools used.*

(10 Points)