**Data Engineering Overview**

There were 6 total tables included in the CSV dataset:

* departments
* dept\_emp
* dept\_manager
* employees
* salaries
* titles

It was discovered that there are two primary tables in the database: "departments" and "employees." The rest of the tables are related to these two main tables. The "departments" table has the primary key "dept\_no," while the "employees" table has the primary key "emp\_no." The other four tables have at least one of these two fields as foreign keys, establishing relationships with the main tables.

**Relationships**

* **departments 1-to-many to dept\_emp**: Employees have the flexibility to be associated with multiple departments over time, such as when they switch between different departments.
* **departments many-to-many to dept\_manager**: Departments can have multiple managers assigned to them at different times. Similarly, managers have the potential to oversee multiple departments over time, at least in theory.
* **employees many-to-many to titles**: Employees have the ability to hold more than one job title due to potential promotions or changes in roles. Conversely, one job title can be assigned to multiple employees, indicating a many-to-many relationship between employees and job titles.
* **employees many-to-many to salaries**: This describes a reciprocal relationship to the previously mentioned scenario. As employees can receive promotions and have multiple job titles, they may correspondingly have more than one salary level. Additionally, multiple employees can have the same salary level, resulting in a many-to-many relationship between employees and salary levels.
* **employees many-to-1 to dept\_emp**: A department has the capacity to accommodate multiple employees.
* **employees many-to-1 to dept\_manager**: A manager has the ability to oversee and lead multiple employees as part of his/her team.

**Data Types**

• All columns related to dates are defined to be of the DATE data type.

• Columns representing variable character data are set to have a maximum length of 100 characters (VARCHAR(100)).

• The employee number (emp\_no) column is set to hold integer values.

• The department number (dept\_no) column is set to hold character values of length 4 (CHAR(4)) since all dept\_no values are consistently in the format 'dXYZ' and do not deviate from a length of 4 characters.

• The gender column is set to hold character values of length 1 (CHAR(1)) since the value can only be either 'M' or 'F'.