**Data Engineering Overview**

There were 6 total tables included in the CSV dataset:

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

After doing analysis and creating ERD diagrams, I conclude that there are to main tables: departments and employees. There other four tables are child tables to either departments or employees tables

**Relationships**

* **departments 1-to-many to dept\_emp**: employees can belong to more than one department over time (for example, switching departments)
* **departments many-to-many to dept\_managers**: departments can multiple managers assigned to them over time. Managers can (theoretically at least) manage multiple departments over time
* **employees many-to-many to titles**: employees can have more than one title as they may be promoted. One title can have many employees assigned with the same title. Therefore, a many-to-many relationship
* **employees many-to-many to salaries**: mirror of previous relationship. Because employees can have more than one title as they get promoted, they may have more than one salary level. One salary level can have many employees as multiple employees can make the same salary. Therefore, a many-to-many relationship
* **employees many-to-1 to dept\_emp**: a department can contain multiple employees
* **employees many-to-1 to dept\_managers**: a manager can have many employees under his/her team

**Data Types**

* All \*date columns defined to be DATE type
* Variable character columns set to VARCHAR
* emp\_no set to INTEGER
* dept\_no set to CHAR(4)
* gender set to CHAR(1) since value is either 'M' or 'F'