Assignment 3:

In this assignment we imported, exported the database and performed indexing on the data.

Firstly, we find a large dataset so that we can see the difference in the execution time.

For importing a .sql file we will first create a database and then import the file using the query below:

>>>mysql -u username -p password database_name < file_name.sql

Query used in the assignment:

>>>mysql -u root spin < load_salaries1.sql (dataset used for Clustering)

>>>mysql -u root sakila < data_sakila.sql (dataset used for Primary Indexing)

Types of indexing:

1. Single valued indexing
    - Primary Indexing
    - Clustering
    - Secondary Indexing
2. Multi valued indexing

For checking the number of records in a table we used the COUNT function.

- Clustering: When we apply indexing on a non-prime attribute it is called clustering.
  Checking for the number of records in the table. This is done by using the "count " function.

```
MariaDB [spin]> select count(*) from salaries;
+-----------+
```

1. Retrieving the salaries whose id is equal to 90000 with and without indexing

   Without Indexing Execution time is 1.881s

```
MariaDB [spin]> select sal from salaries where eid = 90000;
+-------+
| sal   |
+-------+
| 40000 |
| 42979 |
| 43840 |
| 44422 |
| 45251 |
| 48836 |
| 50857 |
| 54562 |
| 58477 |
+-------+
9 rows in set (1.881 sec)
```

If we index the eid attribute which is non-prime the new execution time Is 0.158s. This happens because the pointer will not search in the entire column for these eids sequentially but directly point the eid saving time

```
MariaDB [spin]> select sal from salaries where eid = 90000;
+-------+
| sal   |
+-------+
| 40000 |
| 42979 |
| 43840 |
| 44422 |
| 45251 |
| 48836 |
| 50857 |
| 54562 |
| 58477 |
+-------+
9 rows in set (0.158 sec)
```

2. Retrieving the eids from the table whose salary is equal to 100000 with and without indexing
   In this question we shall index the salary attribute and then execute the queries

   Query for indexing the ‘sal ‘ attribute:

   >>Create index I on salaries(sal);

   With indexing the execution time is 0.066s

   Query is:
   >>select eid from salaries where sal = 100000;

   Without Indexing the execution time is2.218s

   Query is:
   >>select eid from salaries where sal = 100000;