DATA CLEANING

- **Creating a database.**
-> CREATE DATABASE PROJECT;

- **Selecting current Database.**
-> USE PROJECT;

- **Checking the data types of the Table(hr).**
-> DESC HR;

- **Changing the column name from " ï»¿id " to 'emp_id'.**
-> ALTER TABLE HR
   CHANGE COLUMN ï»¿ID EMP_ID VARCHAR(20) NULL;

- **Changing the column(birthdate) format.**
-> UPDATE HR
   SET BIRTHDATE = CASE
   WHEN BIRTHDATE LIKE '%/%' THEN DATE_FORMAT(STR_TO_DATE(BIRTHDATE, '%M/%D/%Y'),
   '%Y-%M-%D')
   WHEN BIRTHDATE LIKE '%-%' THEN DATE_FORMAT(STR_TO_DATE(BIRTHDATE, '%M-%D-%Y'),
   '%Y-%M-%D')
   ELSE NULL
   END;

- **Extracting the birthdate coloumn after format change.**
-> SELECT BIRTHDATE FROM HR;

- **Changing the data type of column(birthdate) to date data type.**
-> ALTER TABLE HR MODIFY COLUMN BIRTHDATE DATE;

- **Changing the column(hiredate) format.**
-> UPDATE HR
   SET HIRE_DATE = CASE
   WHEN HIRE_DATE LIKE '%/%' THEN DATE_FORMAT(STR_TO_DATE(HIRE_DATE, '%M/%D/%Y'),
   '%Y-%M-%D')
   WHEN HIRE_DATE LIKE '%-%' THEN DATE_FORMAT(STR_TO_DATE(HIRE_DATE, '%M-%D-%Y'),
   '%Y-%M-%D')
   ELSE NULL
    END;

- **Extracting the hiredate coloumn after format change.**
-> SELECT HIRE_DATE FROM HR;

- **Changing the data type of column(hiredate) to date data type.**
-> ALTER TABLE HR MODIFY COLUMN HIRE_DATE DATE;

- **Changing the column(termdate) format.**
-> UPDATE HR
   SET TERMDATE = DATE(STR_TO_DATE(TERMDATE, '%Y-%M-%D %H:%I:%S UTC'))
   WHERE TERMDATE IS NOT NULL AND TERMDATE != " ";

- **Extracting the termdate coloumn after format change.**

-> SELECT TERMDATE FROM HR;

- **Updating coloumn(termdate).**

-> UPDATE HR
   SET TERMDATE = NULL
   WHERE TERMDATE = "";

- **Extracting the termdate coloumn after updation.**

-> SELECT TERMDATE FROM HR;

- **Changing the data type of column(termdate) to date data type.**

-> ALTER TABLE HR MODIFY COLUMN TERMDATE DATE;

- **Checking the data type changes.**

-> DESC HR;

- **Checking the column format changes.**

-> SELECT * FROM HR;

- **Adding a new column (Age) and updatiing its values.**

-> ALTER TABLE HR ADD COLUMN AGE INT;
   UPDATE HR SET AGE = TIMESTAMPDIFF(YEAR, BIRTHDATE, CURDATE());

- **Checking the column(age).**

-> SELECT AGE FROM HR;

- **Checking the counts of employees who's termination_date is greater than current date.**

-> SELECT COUNT(*)
    FROM HR
   WHERE TERMDATE > CURDATE();

- **Checking the counts of termdate column where values are "null".**

-> SELECT COUNT(*)
   FROM HR
   WHERE TERMDATE IS NULL;

- **Checking the overall table(hr).**

-> SELECT* FROM HR;