

In [2]:

```
import pandas as pd
import warnings
warnings.filterwarnings("ignore")

#Task1
#Loading the data
def read_data_from_csv():
    #df =read the 'iot_telemetry_data.csv' file
    dtype_options = {'column_name1': str, 'column_name2': str, 'column_name3': str}
    df=pd.read_csv('iot_telemetry_data.csv',dtype=dtype_options, low_memory=False)
    return df

#Task 2: Renaming the Columns
def rename_columns():
    # do not edit the predefined function name
    df=read_data_from_csv()
    columns_rename={'ts':'timestamp','device':'device_id','co':'carbon_monoxide','humidi
                    ', 'light':'light', 'lpg':'liquefied_petroleum_gas', 'motion':'motion', 'smok
    df=df.rename(columns=columns_rename)

    #rename the columns according to the description
    #write your code here
    return df

#Task 3: check for null values
def null_values_check():
    # do not edit the predefined function name
    df=rename_columns()
    #write your code here to check for the null values in the dataset
    null_values=df.isnull().sum()
    return null_values

#Task4 :Removing Duplicates

def remove_duplicates():
    # do not edit the predefined function name
    df=rename_columns()
    #write your code here to drop the duplicates in the dataset and return the dataframe
    df=df.drop_duplicates()

    return df

#Task 5:Handling Missing Values:
def handle_missing_values():
    # do not edit the predefined function name
    df=remove_duplicates()
    #write your code here to drop the missing values in the dataset which returns the da
    df=df.dropna()
    return df

#Task 6:Data Type Conversion:

def convert_data_types():
    # do not edit the predefined function name
    df= handle_missing_values()
    #write your code to change the datatype of each column to respective datatype mentio
```

```
#description and return the dataframe 'df' which contains updated columns datatypes.
df['timestamp'] = df['timestamp'].astype('datetime64[ns]')
df['light']=df['light'].astype('bool')
df['motion']=df['motion'].astype('bool')

return df
```

#Task 7: Export the cleaned dataset to "cleaned_environment.csv"

```
def export_the_dataset():
    # do not edit the predefined function name
    df=convert_data_types()
    df=df.to_csv("cleaned_environment.csv")
    #write your code to export the cleaned dataset and set the index=false and return th
```

#TASK 8: Load the Cleaned dataset 'cleaned_environment.csv' to the database provided.

#follow the instruction in the Task 8 description and complete the task as per it.

#check if mysql table is created using "cleaned_environment"

#Use this final dataset and upload it on the provided database for performing analysis in

#To run this task click on the terminal and click on the run project