```
#!/usr/bin/env python
# coding: utf-8

# In[27]:


import pandas as pd


# In[28]:


data = pd.read_csv("/Users/sathvikreddy/Desktop/
Existing_Buildings_Energy_Performance_Ordinance_Report.csv")


# In[29]:


data = data.dropna(subset = 'Electricity Use - Grid Purchase (kWh)')


# In[30]:


data = data.dropna(subset = 'Natural Gas Use (kBtu)')


# In[32]:


data['Benchmark Status'].unique()


# In[33]:


data = data[data['Benchmark Status'] != '2019 - Violation -
Insufficient Data' ]


# In[34]:


data = data[data['Benchmark Status'] != '2021 - Violation -
Insufficient Data' ]


# In[35]:
```

```
data = data[data['Benchmark Status'] != '2018 - Violation - Did Not
Report' ]


# In[36]:


data = data[data['Benchmark Status'] != '2018 - Violation -
Insufficient Data' ]


# In[37]:


data['Benchmark Status'].unique()


# In[38]:


data


# In[50]:


data.columns


# In[53]:


data = data.rename(columns = {'Building Name':'Building_Name',
                              'Parcel Number':'Parcel_Number',
                              'Building Address':'Building_Adress',
                              'Postal Code':'Postal_Code',
                              'Floor Area':'Floor_Area',
                              'Property Type - Self
Selected':'Property_Type',
                              'PIM Link':'Pim_Link',
                              'Year Built':'Year_Built',
                              'Energy Audit Due Date':'Audit_Duedate',
                              'Energy Audit Status':'Audit_Status',
                              'Benchmark Year':'Benchmark_Year',
                              'ENERGY STAR Score':'Energy_Starscore',
                              'Site EUI (kBtu/
ft2)':'Site_EnergyuseIntensity',
                              'Source EUI (kBtu/
ft2)':'Source_EnergyuseIntensity',
```

```
                              'Percent Better than National Median
Site EUI':'Percent_MedianSite',
                              'Weather Normalized Site EUI (kBtu/
ft2)':'Weather_SiteEnergy',
                              'Weather Normalized Source EUI (kBtu/
ft2)':'Weather_SourceEnergy',
                              'Total GHG Emissions (Metric Tons
CO2e)':'GreenhouseGas_Emission',
                              'Total GHG Emissions Intensity (kGCO2e/
ft2)':'GreenhouseGas_Intensity',
                              'Electricity Use - Grid Purchase
(kWh)':'Electricity_Use',
                              'Natural Gas Use
(kBtu)':'NaturalGas_Use',
                              'District Steam Use (kBtu)':'Steam_Use',
                              'Site Energy Use
(kBtu)':'Site_EnergyUse',
                              'Benchmark Status':'Benchmark_Status',
                              'Reason for Exemption':'Reason'})


# In[54]:


data


# In[57]:


data.to_csv('Modifieddataset.csv')


# In[ ]:




# In[ ]:
```