In [1]:
```
#importing the python libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
#importing the dataset
url = 'http://bit.ly/HDSC-Dataset'
data = pd.read_csv(url)
```

In [3]:
```
data.shape
```

Out[3]:
```
(29523, 11)
```

In [4]:
```
data.head()
```

Out[4]:

	record_id	utility_id_ferc1	report_year	plant_name_ferc1	fuel_type_code_pudl	fuel_unit	fuel_qty_burned	fuel_mmbtu_per_unit	fuel
0	f1_fuel_1994_12_1_0_7	1	1994	rockport	coal	ton	5377489.0	16.590	
1	f1_fuel_1994_12_1_0_10	1	1994	rockport total plant	coal	ton	10486945.0	16.592	
2	f1_fuel_1994_12_2_0_1	2	1994	gorgas	coal	ton	2978683.0	24.130	
3	f1_fuel_1994_12_2_0_7	2	1994	barry	coal	ton	3739484.0	23.950	
4	f1_fuel_1994_12_2_0_10	2	1994	chickasaw	gas	mcf	40533.0	1.000	

In [5]:
```
#Question 1
A = [1,2,3,4,5]
B = [13,21,34]
A_B = A.append(B)
```

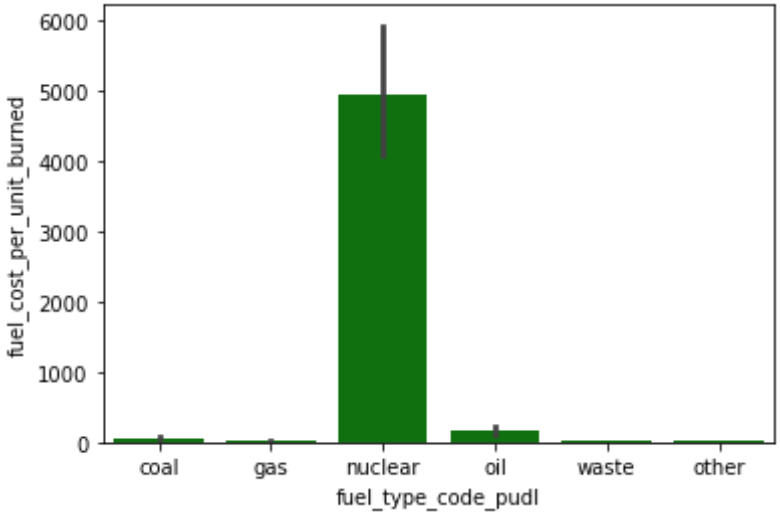
In [6]:
```
#Question 2
np.identity(3)
```

Out[6]:
```
array([[1., 0., 0.],
       [0., 1., 0.],
       [0., 0., 1.]])
```

In [7]:
```
#Question 3
sns.barplot(data['fuel_type_code_pudl'],data['fuel_cost_per_unit_burned'], color = 'green')
```

```
C:\Users\Adeyekun Ahmed\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```

Out[7]:
```
<AxesSubplot:xlabel='fuel_type_code_pudl', ylabel='fuel_cost_per_unit_burned'>
```



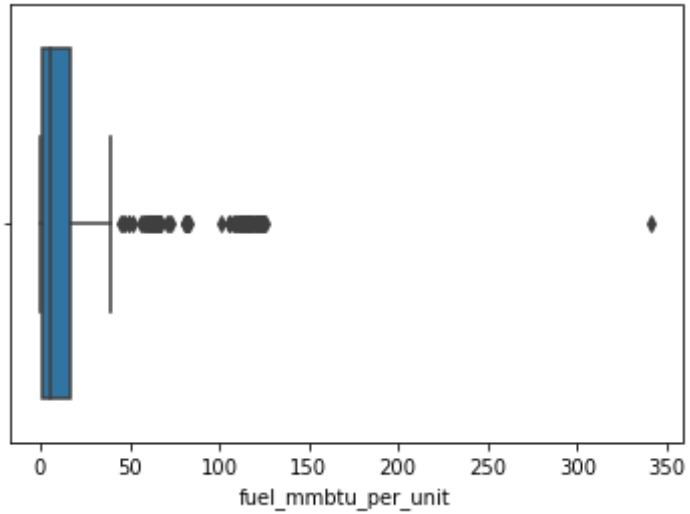
In [8]:
```
#Question 4
# standard deviation
data['fuel_mmbtu_per_unit'].std()
```

Out[8]:
```
10.600220307806886
```

In [11]:
```
sns.boxplot(data['fuel_mmbtu_per_unit'])
```

```
C:\Users\Adeyekun Ahmed\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```

Out[11]:
```
<AxesSubplot:xlabel='fuel_mmbtu_per_unit'>
```

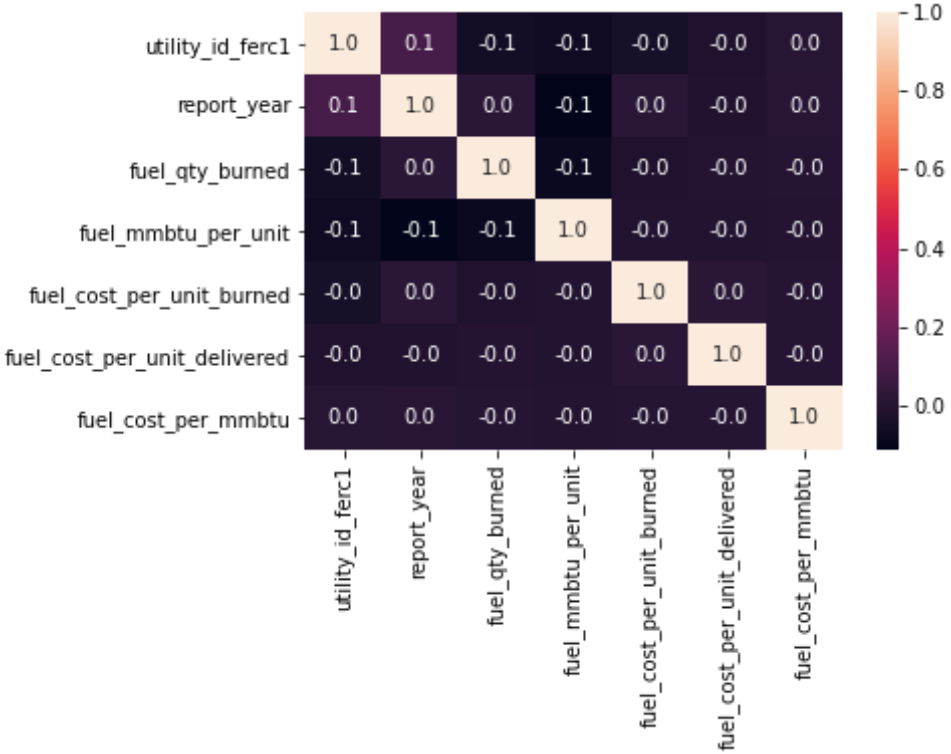


In [12]:
```
#Question 6
# The total and missing values
data.isnull().sum()
```

Out[12]:
```
record_id                       0
utility_id_ferc1                0
report_year                     0
plant_name_ferc1                0
fuel_type_code_pudl             0
fuel_unit                     180
fuel_qty_burned                 0
fuel_mmbtu_per_unit             0
fuel_cost_per_unit_burned       0
fuel_cost_per_unit_delivered    0
fuel_cost_per_mmbtu             0
dtype: int64
```

In [13]:
```
#Question 7
# Categorical and mode imputation
```

In [26]:
```
#Question 8
sns.heatmap(data.corr(), annot = True , fmt = '0.1f')
plt.show()
```

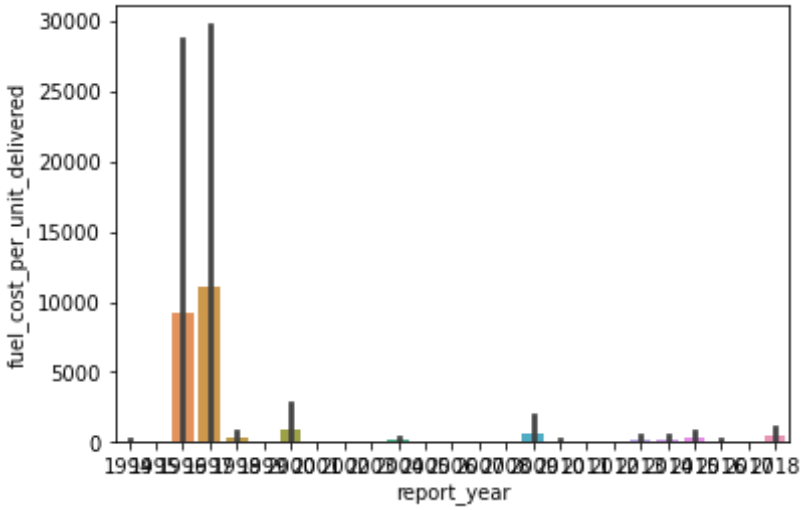


In [41]:
```
#Question 10

sns.barplot(data['report_year'], data['fuel_cost_per_unit_delivered'])
plt.show()
```

```
C:\Users\Adeyekun Ahmed\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following vari
```

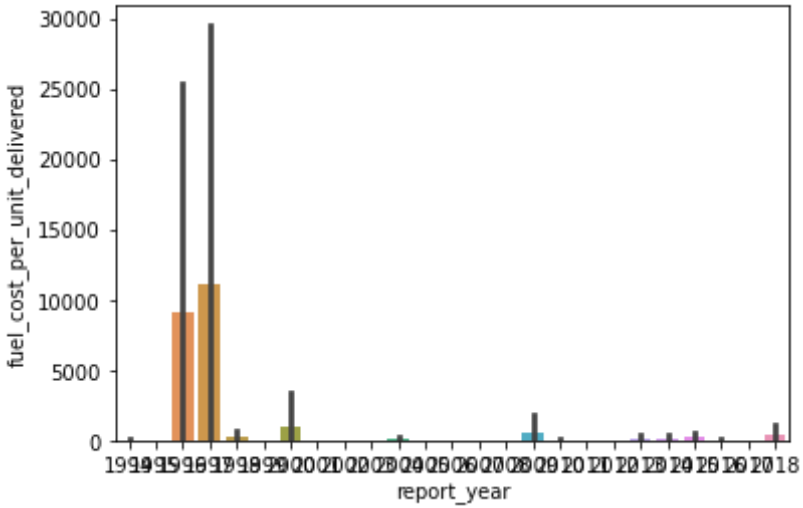
```
ables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other
arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```



In [45]:
```
sns.barplot(data['report_year'], data['fuel_cost_per_unit_delivered'])
```

```
C:\Users\Adeyekun Ahmed\anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following vari
ables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other
arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```

Out[45]: <AxesSubplot:xlabel='report_year', ylabel='fuel_cost_per_unit_delivered'>



In [ ]:

In [ ]: