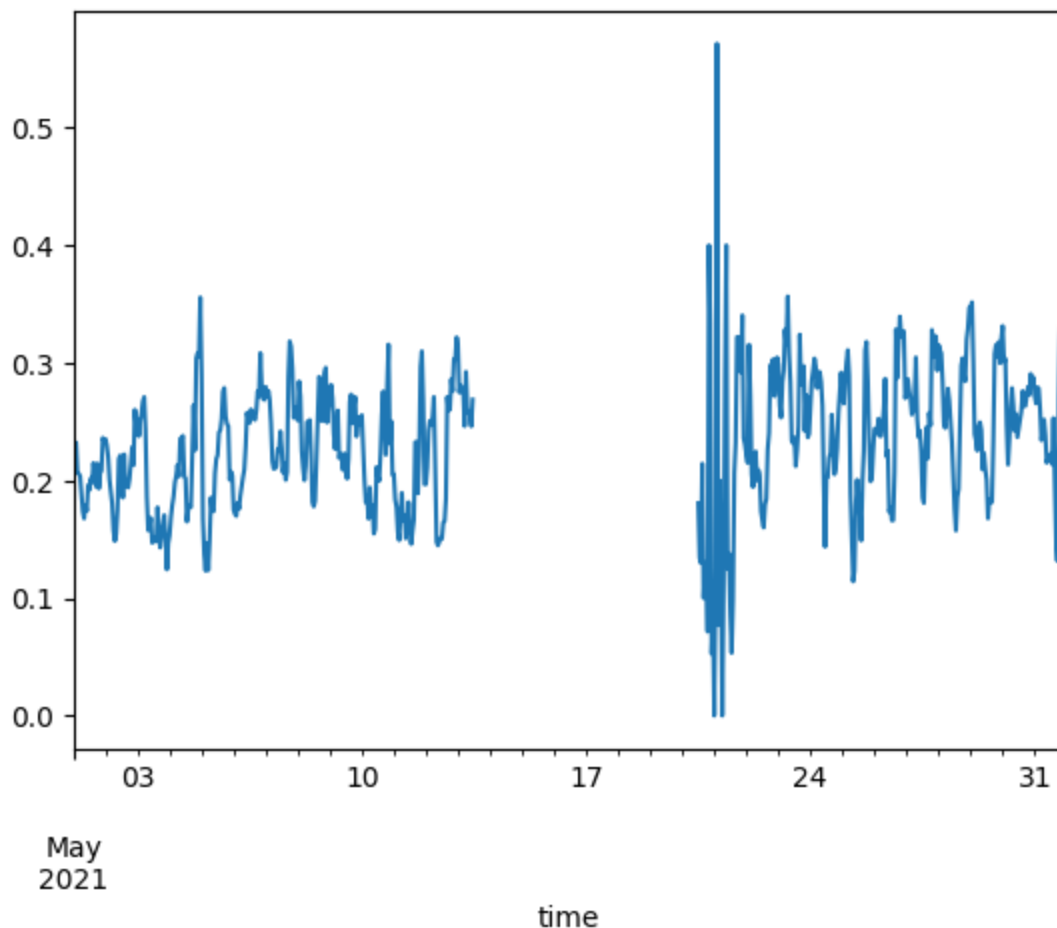


```
In [1]: # A2Q4
# independently coded by Shao Shi
# this is a single particle BC number count data preprocessed in matlab
# the original data is a true data measured at SUSTech Lab 930 on the instrument Single Particle
```

```
In [2]: # 1
# firstly, time of no sample value is converted to NaN value
import pandas as pd
from matplotlib import pyplot as plt

raw = pd.read_csv('202105_withTableHeader.csv',
                  parse_dates=['time'],
                  na_values={'BC/total': 'NaN', 'fresh/total': 'NaN', 'aged/total': 'NaN'})
```

```
In [3]: # 2
# plot the time series of the variable BC fraction
raw = raw.set_index("time")
raw['BC/total'].plot()
plt.show()
```



```
In [4]: # 3
# some simple statistical analysis are going to be computed
print(raw['BC/total'].mean())
print(raw['BC/total'].max())
print(raw['BC/total'].min())
print(raw['BC/total'].std())
print(raw['BC/total'].var())
raw['BC/total'].hist()
plt.show()
# one may find that:
```

```
# the maximum value of BC fraction can reach 0.57, this is a result of special pollution case at  
# the mean value of BC fraction in May 2021 is 0.23, which is contributed by shipping event and  
# min value is 0, due to the fresh air condition in some days  
# the standard variation is 0.058, meaning that the BC fraction is relatively stable  
# one can find the distribution of BC fraction is close to normal distribution
```

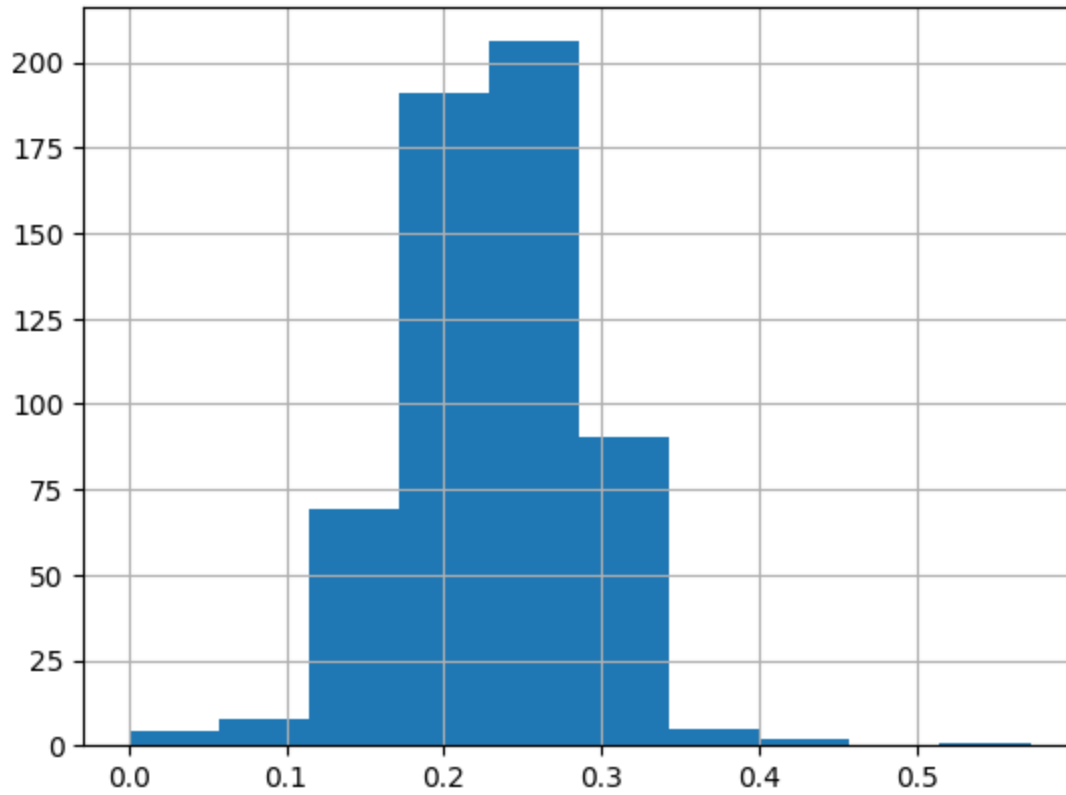
```
0.23166923685416663
```

```
0.571428571
```

```
0.0
```

```
0.0582121734539424
```

```
0.003388657138231876
```



In [ ]: