

```
In [1]: #pandas module is needed to read the csv for this quiz
import pandas as pd
import numpy as np
import matplotlib as plt

#reading the csv file using pd.read_csv
rt = pd.read_csv('RT_IOT2022.csv')

#outputting the dataframe
rt.head()
```

```
Out[1]:
```

	no	id.orig_p	id.resp_p	proto	service	flow_duration	fwd_pkts_tot	bwd_pkts_tot	fwd
0	0	38667	1883	tcp	mqtt	32.011598	9	5	
1	1	51143	1883	tcp	mqtt	31.883584	9	5	
2	2	44761	1883	tcp	mqtt	32.124053	9	5	
3	3	60893	1883	tcp	mqtt	31.961063	9	5	
4	4	51087	1883	tcp	mqtt	31.902362	9	5	

5 rows × 85 columns



```
In [2]: flow_mean = np.mean(rt['flow_duration'])
# this shows the average flow duration
print('The average flow duration is', flow_mean)
```

The average flow duration is 3.8095657699992693

```
In [3]: rt.isnull().sum() #checking for null entries in the data
```

```
Out[3]: no                                0
id.orig_p                               0
id.resp_p                               0
proto                                   0
service                                0
..
idle.std                               0
fwd_init_window_size                   0
bwd_init_window_size                   0
fwd_last_window_size                   0
Attack_type                            0
Length: 85, dtype: int64
```

```
In [4]: rt.dtypes #displaying the data types of each column
```

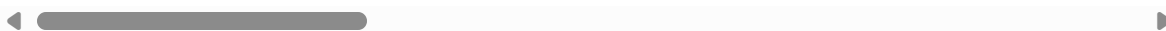
```
Out[4]: no                                int64
id.orig_p                                int64
id.resp_p                                int64
proto                                    object
service                                  object
...
idle.std                                float64
fwd_init_window_size                    int64
bwd_init_window_size                    int64
fwd_last_window_size                    int64
Attack_type                             object
Length: 85, dtype: object
```

```
In [5]: rt = rt.dropna() # removing null entries in the dataframe
rt
```

```
Out[5]:
```

	no	id.orig_p	id.resp_p	proto	service	flow_duration	fwd_pkts_tot	bwd_pkts_1
0	0	38667	1883	tcp	mqtt	32.011598	9	
1	1	51143	1883	tcp	mqtt	31.883584	9	
2	2	44761	1883	tcp	mqtt	32.124053	9	
3	3	60893	1883	tcp	mqtt	31.961063	9	
4	4	51087	1883	tcp	mqtt	31.902362	9	
...
123112	2005	59247	63331	tcp	-	0.000006	1	
123113	2006	59247	64623	tcp	-	0.000007	1	
123114	2007	59247	64680	tcp	-	0.000006	1	
123115	2008	59247	65000	tcp	-	0.000006	1	
123116	2009	59247	65129	tcp	-	0.000006	1	

123117 rows × 85 columns



In []: