

df - Excel

source\_ip,destination\_ip,start\_time,source\_port,destination\_port,flags,site,asn,num\_packets,num\_bytes

source_ip	destination_ip	start_time	source_port	destination_port	flags	site	asn	num_packets	num_bytes
135.b1d10.d1c38.20	135.0777d.04511.237	1415749946	22	45092	-AP---	45c48	c4ca4	6	504
135.0777d.04511.237	135.b1d10.13fe9.91	1415477729	45603	22	-A----,-AP---	45c48	c4ca4	11	956
135.0777d.04511.237	135.b1d10.13fe9.71	1415702327	45596	22	-A----,-AP---	45c48	c4ca4	11	956
135.b1d10.d1c38.119	135.0777d.04511.237	1415754478	22	50107	-AP---	45c48	c4ca4	6	504
135.b1d10.13fe9.56	135.0777d.04511.237	1415749597	22	45580	-AP---	45c48	c4ca4	6	504
135.b1d10.d1c38.72	135.0777d.04511.237	1415745508	22	51595	-AP---	45c48	c4ca4	6	504
135.b1d10.d1c38.126	135.0777d.04511.237	1415754317	22	38662	-AP---	45c48	c4ca4	6	504
135.0777d.04511.237	135.b1d10.d1c38.37	1415427333	48823	22	-A----,-AP---	45c48	c4ca4	12	1008
135.0777d.04511.237	135.b1d10.d1c38.112	1415584036	51452	22	-A----,-AP---	45c48	c4ca4	12	1008
135.0777d.04511.237	135.b1d10.d1c38.20	1415582789	45092	22	-A----,-AP---	45c48	c4ca4	12	1008
135.b1d10.13fe9.91	135.0777d.04511.237	1415747239	22	45603	-AP---	45c48	c4ca4	6	504
135.0777d.04511.237	135.b1d10.13fe9.85	1415678603	59307	22	-A----,-AP---	45c48	c4ca4	12	1008
135.b1d10.13fe9.118	135.0777d.04511.237	1415743957	22	36831	-AP---	45c48	c4ca4	6	504
135.b1d10.d1c38.79	135.0777d.04511.237	1415754477	22	51495	-AP---	45c48	c4ca4	6	504
135.b1d10.13fe9.64	135.0777d.04511.237	1415754296	22	59296	-AP---	45c48	c4ca4	6	504
135.0777d.04511.237	135.b1d10.d1c38.153	1415422624	38605	22	-A----,-AP---	45c48	c4ca4	12	1008
135.0777d.04511.237	135.b1d10.d1c38.208	1415477735	51461	22	-A----,-AP---	45c48	c4ca4	12	1008
135.0777d.04511.237	135.b1d10.d1c38.225	1415753589	51457	22	-A----,-AP---	45c48	c4ca4	12	1008
135.0777d.04511.237	135.b1d10.9cfd9.109	1415753380	41070	22	-A----,-AP---	45c48	c4ca4	12	1008
135.0777d.04511.237	135.b1d10.9cfd9.109	1415668378	33003	22	-A----,-AP---	45c48	c4ca4	12	1008
135.b1d10.9cfd9.73	135.0777d.04511.237	1415745478	22	38660	-A----,-AP---	45c48	c4ca4	6	504
135.b1d10.9cfd9.100	135.0777d.04511.237	1415753850	22	32981	-A----,-AP---	45c48	c4ca4	6	504
135.0777d.04511.237	135.b1d10.9cfd9.100	1415754422	32981	22	-A----,-AP---	45c48	c4ca4	12	1008
135.b1d10.9cfd9.56	135.0777d.04511.237	1415742950	22	38594	-A----,-AP---	45c48	c4ca4	6	504

## Данные

localhost:8889/notebooks/PROJECT1.ipynb#We-want-to-be-able-to-predict-Y-which-is-our-price-variable.

jupyter PROJECT1 Last Checkpoint: час назад (autosaved)

File Edit View Insert Cell Kernel Widgets Help

### DATA ANALYSIS

```
In [1]: import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

```
In [2]: my_project = pd.read_csv(r'C:\Users\user\Desktop\df.csv')
my_project.head()
```

```
Out[2]:
```

	source_ip	destination_ip	start_time	source_port	destination_port	flags	site	asn	num_packets	num_bytes
0	135.b1d10.d1c38.20	135.0777d.04511.237	1415749946	22	45092	-AP---	45c48	c4ca4	6	504
1	135.0777d.04511.237	135.b1d10.13fe9.91	1415477729	45603	22	-A----,-AP---	45c48	c4ca4	11	956
2	135.0777d.04511.237	135.b1d10.13fe9.71	1415702327	45596	22	-A----,-AP---	45c48	c4ca4	11	956
3	135.b1d10.d1c38.119	135.0777d.04511.237	1415754478	22	50107	-AP---	45c48	c4ca4	6	504
4	135.b1d10.13fe9.56	135.0777d.04511.237	1415749597	22	45580	-AP---	45c48	c4ca4	6	504

```
In [3]: my_project.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 337940 entries, 0 to 337939
Data columns (total 10 columns):
source_ip      337940 non-null object
destination_ip 337940 non-null object
start_time     337940 non-null object
```

Я начала, но не все оказалось неправильным и я забросила