

```
1 import pandas as pd
2 import numpy as np
3 from scipy.stats.mstats import hmean
4
5 data = pd.read_csv('/Users/artiomkholodkov/
Downloads/1016/1016.csv')
6 data.info()
7 needed_data = data.loc[:,['3','4','5','6','7',
,'14']]
8 needed_data.columns = ['DATE', 'TIME', 'LANE'
, 'DIRECTION', 'SPEED', 'HEADWAY']
9
10 date = pd.Series(needed_data['DATE'])
11 modified_date = date.apply(lambda x: x[:8].
lstrip().replace("/", "-"))
12
13 time = pd.Series(needed_data['TIME'])
14 modified_time = time.apply(lambda x: x[-8:].
lstrip()) #with spaces
15
16 new_time = modified_date.combine(
modified_time, (lambda x1, x2: x1 + ' ' + x2)
)
17
18 needed_data_2 = data.loc[:,['5','6','7','14']
]
19 data_frame = pd.concat([new_time,
needed_data_2], axis=1)
20 data_frame.columns = ['TIME', 'LANE', '
DIRECTION', 'SPEED', 'HEADWAY']
21 print(data_frame)
22 data_frame.to_csv('cleaned_data_1016.csv')
23
```