

- Pandas: 导入数据

```
1 import pandas as pd
2
3 names = ['mpg', 'cylinders', 'displacement', 'horsepower',
4          'weight', 'acceleration', 'model year', 'origin', 'car name']
5
6 df = pd.read_csv('https://archive.ics.uci.edu/ml/machine-learning-databases/'+
7                  'auto-mpg/auto-mpg.data',
8                  header=None, delim_whitespace=True, names=names, na_values='?')
9 # 查看前六行
10 df.head(6)
11
12 # 一些属性
13 # 返回 (行数, 列数)
14 df.shape
15 # 返回列名
16 df.columns
17 df.columns.tolist()
18 # 返回行数索引
19 df.index
20 # 返回每行的数据
21 df.values
22
23 # 选取其中某些列
24 df2 = df[['cylinders', 'horsepower']]
25 # 转化成数组
26 y = np.array(df['mpg'])
27 # 转化成多维数组, 注意这里要用两层 [[]] 包起来! 不知道为啥
28 y = np.array(df[['mpg', 'xxx']])
```