# Python Data Analysis

**To read data from a file called "my\_data.csv",** we first import the csv module:

import csv

Next, we open the file:

f = open("my\_data.csv")

Then, we call the module's reader() function:

csvreader = csv.reader(f)

Finally, we convert the result to a list:

my\_data = list(csvreader)

## 6: Counting How Many Times A Team Won

1. 表格list有row 有column，循环输出的时候可以：

for row in nfl:

1. 赢的球队一般名字在前。

## [Data Analysis with Pandas: Intermediate](https://www.dataquest.io/course/data-analysis-intermediate) (Pandas)

1. Import **pandas**
2. Read.csv()
3. Head() : When you call the head() method, pandas will return a new dataframe containing just the first 5 rows:  
    first\_rows = food\_info.head()  
    print(food\_info.head(3)) # First 3 rows.
4. Columns(): To access **the full list** of column names, use the columns attribute
   1. column\_names = food\_info.columns
5. Shape():
   1. # Returns the tuple (8618,36) and assigns to `dimensions`.

dimensions = food\_info.shape

# The number of rows, 8618.

num\_rows = dimensions[0]

# The number of columns, 36.

num\_cols = dimensions[1]

1. loc[]: The loc[] method allows you to select rows by row labels. Eg: choose the 100th row🡺 loc[99]
2. 如果求最后几行，先用shape[0]求出总长度，再减去所求长度，得到需求的第一行。
3. Column select:
   1. # Series object representing the "NDB\_No" column.

ndb\_col = food\_info["NDB\_No"]

# You can instead access a column by passing in a string variable.

col\_name = "NDB\_No"

ndb\_col = food\_info[col\_name]

1. Tolist(): food\_info.columns.tolist()
2. 赋值的时候：gram\_df = food\_info[df]，其中df为选中列的名字list， food\_info[df]是全部列中的选中列。
3. Endswith(“\*\*\*”): list.endswith(“\*\*\*”)
4. Eg
   1. import pandas
   2. food\_info = pandas.read\_csv("food\_info.csv")
   3. gram\_columns = food\_info.columns.tolist()
   4. df = []
   5. for i in gram\_columns:
   6. if i.endswith("(g)"): ### 找以（g）结尾的title
   7. df += [i]
   8. gram\_df = food\_info[df]