In [2]:

```
#1. create a sample JSON object from titanic train dataset

# https://raw.githubusercontent.com/pcsanwald/kaggle-titanic/master/train.csv
```

In [3]:

```
# Read the JSON object and convert it to pandas data frame
```

In [4]:

```
# match the top 5 records from JSON read and actual titanic dataset
```

In [5]:

```
#2. read the HTML file as given below

# http://vincentarelbundock.github.io/Rdatasets/datasets.html
```

In [6]:

```
# create a data frame from the above object
```

In [7]:

```
## Create URL to JSON file (alternatively this can be a filepath)
url = 'https://raw.githubusercontent.com/chrisalbon/simulated_datasets/master/data.json'
```

In [8]:

```
# Load library
import pandas as pd
```

In [9]:

```python
# Load the first sheet of the JSON file into a data frame
df = pd.read_json(url, orient='columns')

# View the first ten rows
df.head(10)
```

Out[9]:

	category	datetime	integer
0	0	2015-01-01 00:00:00	5
1	0	2015-01-01 00:00:01	5
10	0	2015-01-01 00:00:10	5
11	0	2015-01-01 00:00:11	5
12	0	2015-01-01 00:00:12	8
13	0	2015-01-01 00:00:13	9
14	0	2015-01-01 00:00:14	8
15	0	2015-01-01 00:00:15	8
16	0	2015-01-01 00:00:16	2
17	0	2015-01-01 00:00:17	1

In [ ]: