pandas_data_manipulation_workbook_answers

September 22, 2020

1   PANDAS DATA MANIPULATION WORKBOOK ANSWER

Remember, there are different ways to write code to get the same answer, so your answer can be correct and different to the answer example!

If you feel stuck and want some in person help, then have a look at the events page to join in a workshop https://swamphen.co.uk/events.

```python
In [ ]: # import pandas
        import pandas as pd
```

```python
In [ ]: # import data directly from web link
        # https://vincentarelbundock.github.io/Rdatasets/csv/carData/Greene.csv
        # information at https://vincentarelbundock.github.io/Rdatasets/doc/carData/Greene.htm
        greene = pd.read_csv('https://vincentarelbundock.github.io/Rdatasets/csv/carData/Greene
```

```python
In [ ]: # print out the head
        greene.head()
```

```python
In [ ]: # print rows 8 to 12
        greene[8:12]
```

```python
In [ ]: # print out the last column only
        greene['success']
```

```python
In [ ]: # select columns language and location, rows 10 to 15
        greene[['language', 'location']][10:15]
```

```python
In [ ]: # pick a location from the list above and select all the data at this location
        toronto = greene['location'] == 'Toronto'
```

```python
In [ ]: # use this to filter the whole pandas data frame
        greene[toronto]
```

```python
In [ ]: # how many cases were in your chosen city?
        len(greene[toronto])
```

```python
In [ ]: # sort the entire pandas data frame on language
        greene.sort_values('language')
```

```
In [ ]: # sort the pandas data frame on language then location
        greene.sort_values(['language', 'location'])

In [ ]: # are there locations that speak more than one language?

        # yes montreal

In [ ]: # find out the rows that hav Somalia in them
        greene[greene['nation'].isin(['Somalia'])]
```