Python_Module9

September 27, 2020

0.1 Author: Joseph Vargovich

[2]:
```
#Import libraries
import pandas as pd
```

1 Exercise 1: Reading in an excel file with skipped lines.

[4]:
```
#Pandas does this all quite well. Use the pandas.read_excel() function
data1 = pd.read_excel("Example_5.xls", "RawData", usecols="A:C", skiprows=4)

data1.tail() #The final element is 30 because the first row is numbered as 0.
```

[4]:

	Girth	Height	Volume
26	17.5	82	55.7
27	17.9	80	58.3
28	18.0	80	51.5
29	18.0	80	51.0
30	20.6	87	77.0

2 Exercise 2 - Reading in another excel spreadsheet and manipulating it.

[19]:
```
#Load the data from excel file
data2 = pd.read_excel("Example_3.xls", "data")

#Drop rows and columns with all NA values.
data2 = data2[data2['model'].notna()]
data2 = data2.dropna(axis = 1, how = 'all')

#Drop a lingering NA column.
data2 = data2.drop(['Unnamed: 19'], axis=1)
data2.tail()
```

[19]:

	model	mpg	cyl	disp	hp	drat	wt	qsec	vs	\
28	Ford Pantera L	15.8	8.0	351	264.0	4.22	3.170	14.50	0.0	

```
29        Ferrari Dino  19.7  6.0  145  175.0     3.62  2.770  15.50  0.0
30       Maserati Bora  15.0  8.0  301  335.0     3.54  3.570  14.60  0.0
31          Volvo 142E  21.4  4.0  121  109.0     4.11  2.780  18.60  1.0
32  Tesla ModelS P100D  98.0  NaN   NA  778.0 -9999.00  4.941  10.41  NaN

     am  gear  carb
28  1.0   5.0   4.0
29  1.0   5.0   6.0
30  1.0   5.0   8.0
31  1.0   4.0   2.0
32  0.0   1.0   NaN
```