Ex1_Win_SASPy_etc

April 25, 2019

```
In [56]: import pandas as pd
         df=pd.read_csv('/SASCourse/Week14/TV_Data_noheader.csv', names=['opinion', 'party', 'in
         print(df)
```

```
              opinion        party  income  age
0               Agree   Republican   35000   30
1            Disagree     Democrat   40000   50
2      Strongly Agree  Independent  100000   40
3          No Opinion  Green Party   90000   45
4   Strongly Disagree     Democrat   65000   45
5            Disagree  Independent   55000   39
6      Strongly Agree  Green Party   75000   80
7          No Opinion  Independent   55000   70
8   Strongly Disagree  Green Party   28000   30
9          No Opinion     Democrat   80000   45
10  Strongly Disagree  Independent   45000   25
11           Disagree  Green Party  110000   45
12              Agree     Democrat   55000   60
13           Disagree  Independent   37000   29
14     Strongly Agree  Green Party   81000   81
15         No Opinion  Green Party   59000   40
16  Strongly Disagree     Democrat  150000   45
17           Disagree  Independent   67000   61
18     Strongly Agree  Green Party   78000   45
19         No Opinion  Independent  200000   52
20  Strongly Disagree  Green Party    8000   21
21         No Opinion  Green Party   55000   50
22  Strongly Disagree  Independent   70000   39
23           Disagree  Green Party   67000   40
```

```
In [57]: import pandas as pd
         url = "https://simple.wikipedia.org/wiki/List_of_U.S._states"
         mylist = pd.read_html(url)
         print(mylist)
```

```
[                          0          1              2              3   \
0                    Name &        Abbr         Cities     Established
```

1		Capital	Largest[3]	mi2	km2
2	Alabama	AL	Montgomery	Birmingham	
3	Alaska	AK	Juneau	Anchorage	
4	Arizona	AZ	Phoenix	Feb 14, 1912	
5	Arkansas	AR	Little Rock	Jun 15, 1836	
6	California	CA	Sacramento	Los Angeles	
7	Colorado	CO	Denver	Aug 1, 1876	
8	Connecticut	CT	Hartford	Bridgeport	
9	Delaware	DE	Dover	Wilmington	
10	Florida	FL	Tallahassee	Jacksonville	
11	Georgia	GA	Atlanta	Jan 2, 1788	
12	Hawaii	HI	Honolulu	Aug 21, 1959	
13	Idaho	ID	Boise	Jul 3, 1890	
14	Illinois	IL	Springfield	Chicago	
15	Indiana	IN	Indianapolis	Dec 11, 1816	
16	Iowa	IA	Des Moines	Dec 28, 1846	
17	Kansas	KS	Topeka	Wichita	
18	Kentucky[upper-alpha 2]	KY	Frankfort	Louisville	
19	Louisiana	LA	Baton Rouge	New Orleans	
20	Maine	ME	Augusta	Portland	
21	Maryland	MD	Annapolis	Baltimore	
22	Massachusetts[upper-alpha 2]	MA	Boston	Feb 6, 1788	
23	Michigan	MI	Lansing	Detroit	
24	Minnesota	MN	St. Paul	Minneapolis	
25	Mississippi	MS	Jackson	Dec 10, 1817	
26	Missouri	MO	Jefferson City	Kansas City	
27	Montana	MT	Helena	Billings	
28	Nebraska	NE	Lincoln	Omaha	
29	Nevada	NV	Carson City	Las Vegas	
30	New Hampshire	NH	Concord	Manchester	
31	New Jersey	NJ	Trenton	Newark	
32	New Mexico	NM	Santa Fe	Albuquerque	
33	New York	NY	Albany	New York	
34	North Carolina	NC	Raleigh	Charlotte	
35	North Dakota	ND	Bismarck	Fargo	
36	Ohio	OH	Columbus	Mar 1, 1803	
37	Oklahoma	OK	Oklahoma City	Nov 16, 1907	
38	Oregon	OR	Salem	Portland	
39	Pennsylvania[upper-alpha 2]	PA	Harrisburg	Philadelphia	
40	Rhode Island[upper-alpha 3]	RI	Providence	May 29, 1790	
41	South Carolina	SC	Columbia	Charleston	
42	South Dakota	SD	Pierre	Sioux Falls	
43	Tennessee	TN	Nashville	Jun 1, 1796	
44	Texas	TX	Austin	Houston	
45	Utah	UT	Salt Lake City	Jan 4, 1896	
46	Vermont	VT	Montpelier	Burlington	
47	Virginia[upper-alpha 2]	VA	Richmond	Virginia Beach	
48	Washington	WA	Olympia	Seattle	

49	West Virginia	WV	Charleston	Jun 20, 1863	
50	Wisconsin	WI	Madison	Milwaukee	
51	Wyoming	WY	Cheyenne	Jul 10, 1890	

	4	5	6	7	\
0	Population[upper-alpha 1][1]	Total area[2]	Land area[2]	Water area[2]	
1	mi2	km2	mi2	km2	
2	Dec 14, 1819	4874747	52420	135767	
3	Jan 3, 1959	739795	665384	1723337	
4	7016270	113990	295234	113594	
5	3004279	53179	137732	52035	
6	Sep 9, 1850	39536653	163695	423967	
7	5607154	104094	269601	103642	
8	Jan 9, 1788	3588184	5543	14357	
9	Dec 7, 1787	961939	2489	6446	
10	Mar 3, 1845	20984400	65758	170312	
11	10429379	59425	153910	57513	
12	1427538	10932	28313	6423	
13	1716943	83569	216443	82643	
14	Dec 3, 1818	12802023	57914	149995	
15	6666818	36420	94326	35826	
16	3145711	56273	145746	55857	
17	Jan 29, 1861	2913123	82278	213100	
18	Jun 1, 1792	4454189	40408	104656	
19	Apr 30, 1812	4684333	52378	135659	
20	Mar 15, 1820	1335907	35380	91633	
21	Apr 28, 1788	6052177	12406	32131	
22	6859819	10554	27336	7800	
23	Jan 26, 1837	9962311	96714	250487	
24	May 11, 1858	5576606	86936	225163	
25	2984100	48432	125438	46923	
26	Aug 10, 1821	6113532	69707	180540	
27	Nov 8, 1889	1050493	147040	380831	
28	Mar 1, 1867	1920076	77348	200330	
29	Oct 31, 1864	2998039	110572	286380	
30	Jun 21, 1788	1342795	9349	24214	
31	Dec 18, 1787	9005644	8723	22591	
32	Jan 6, 1912	2088070	121590	314917	
33	Jul 26, 1788	19849399	54555	141297	
34	Nov 21, 1789	10273419	53819	139391	
35	Nov 2, 1889	755393	70698	183108	
36	11658609	44826	116098	40861	
37	3930864	69899	181037	68595	
38	Feb 14, 1859	4142776	98379	254799	
39	Dec 12, 1787	12805537	46054	119280	
40	1059639	1545	4001	1034	
41	May 23, 1788	5024369	32020	82933	
42	Nov 2, 1889	869666	77116	199729	

43	6715984	42144	109153	41235
44	Dec 29, 1845	28304596	268596	695662
45	3101833	84897	219882	82170
46	Mar 4, 1791	623657	9616	24906
47	Jun 25, 1788	8470020	42775	110787
48	Nov 11, 1889	7405743	71298	184661
49	1815857	24230	62756	24038
50	May 29, 1848	5795483	65496	169635
51	579315	97813	253335	97093

	8	9	10	11	12
0	Numberof Reps.	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN	NaN
2	50645	131171.0	1775.0	4597.0	7.0
3	570641	1477953.0	94743.0	245384.0	1.0
4	294207	396.0	1026.0	9.0	NaN
5	134771	1143.0	2961.0	4.0	NaN
6	155779	403466.0	7916.0	20501.0	53.0
7	268431	452.0	1170.0	7.0	NaN
8	4842	12542.0	701.0	1816.0	5.0
9	1949	5047.0	540.0	1399.0	1.0
10	53625	138887.0	12133.0	31424.0	27.0
11	148959	1912.0	4951.0	14.0	NaN
12	16635	4509.0	11678.0	2.0	NaN
13	214045	926.0	2398.0	2.0	NaN
14	55519	143793.0	2395.0	6202.0	18.0
15	92789	593.0	1537.0	9.0	NaN
16	144669	416.0	1077.0	4.0	NaN
17	81759	211754.0	520.0	1346.0	4.0
18	39486	102269.0	921.0	2387.0	6.0
19	43204	111898.0	9174.0	23761.0	6.0
20	30843	79883.0	4537.0	11750.0	2.0
21	9707	25142.0	2699.0	6990.0	8.0
22	20202	2754.0	7134.0	9.0	NaN
23	56539	146435.0	40175.0	104052.0	14.0
24	79627	206232.0	7309.0	18930.0	8.0
25	121531	1508.0	3907.0	4.0	NaN
26	68742	178040.0	965.0	2501.0	8.0
27	145546	376962.0	1494.0	3869.0	1.0
28	76824	198974.0	524.0	1356.0	3.0
29	109781	284332.0	791.0	2048.0	4.0
30	8953	23187.0	397.0	1027.0	2.0
31	7354	19047.0	1368.0	3544.0	12.0
32	121298	314161.0	292.0	757.0	3.0
33	47126	122057.0	7429.0	19240.0	27.0
34	48618	125920.0	5201.0	13471.0	13.0
35	69001	178711.0	1698.0	4397.0	1.0
36	105829	3965.0	10269.0	16.0	NaN

```
37          177660     1304.0   3377.0       5.0   NaN
38           95988   248608.0   2391.0    6191.0   5.0
39           44743   115883.0   1312.0    3397.0  18.0
40            2678      511.0   1324.0       2.0   NaN
41           30061    77857.0   1960.0    5076.0   7.0
42           75811   196350.0   1305.0    3379.0   1.0
43          106798      909.0   2355.0       9.0   NaN
44          261232   676587.0   7365.0   19075.0  36.0
45          212818     2727.0   7064.0       4.0   NaN
46            9217    23871.0    400.0    1035.0   1.0
47           39490   102279.0   3285.0    8508.0  11.0
48           66456   172119.0   4842.0   12542.0  10.0
49           62259      192.0    497.0       3.0   NaN
50           54158   140268.0  11339.0   29367.0   8.0
51          251470      720.0   1864.0       1.0   NaN  ]
```

```
In [58]: import saspy
         import pandas as pd
         sas = saspy.SASsession(cfgname='winlocal')
         cars = sas.sasdata("CARS","SASHELP")
         cars.describe()
```

```
SAS Connection established. Subprocess id is 13184
```

```
Out[58]:       Variable            Label    N  NMiss   Median          Mean  \
         0         MSRP              NaN  428      0  27635.0  32774.855140
         1      Invoice              NaN  428      0  25294.5  30014.700935
         2   EngineSize  Engine Size (L)  428      0      3.0      3.196729
         3    Cylinders              NaN  426      2      6.0      5.807512
         4   Horsepower              NaN  428      0    210.0    215.885514
         5    MPG_City       MPG (City)  428      0     19.0     20.060748
         6  MPG_Highway    MPG (Highway)  428      0     26.0     26.843458
         7       Weight     Weight (LBS)  428      0   3474.5   3577.953271
         8    Wheelbase   Wheelbase (IN)  428      0    107.0    108.154206
         9       Length      Length (IN)  428      0    187.0    186.362150

                  StdDev      Min       P25      P50      P75       Max
         0  19431.716674  10280.0  20329.50  27635.0  39215.0  192465.0
         1  17642.117750   9875.0  18851.00  25294.5  35732.5  173560.0
         2      1.108595      1.3      2.35      3.0      3.9       8.3
         3      1.558443      3.0      4.00      6.0      6.0      12.0
         4     71.836032     73.0    165.00    210.0    255.0     500.0
         5      5.238218     10.0     17.00     19.0     21.5      60.0
         6      5.741201     12.0     24.00     26.0     29.0      66.0
         7    758.983215   1850.0   3103.00   3474.5   3978.5    7190.0
```

```
8      8.311813     89.0    103.00    107.0    112.0     144.0
9     14.357991    143.0    178.00    187.0    194.0     238.0
```

```
In [59]: import saspy
         import pandas as pd
         sas = saspy.SASsession(cfgname='winlocal')
         w_class = sas.sasdata("CLASS","SASHELP")
         w_class.describe()
```

```
SAS Connection established. Subprocess id is 12908
```

```
Out[59]:   Variable   N  NMiss  Median        Mean     StdDev   Min   P25   P50    P75  \
         0      Age  19      0    13.0   13.315789   1.492672  11.0  12.0  13.0   15.0
         1   Height  19      0    62.8   62.336842   5.127075  51.3  57.5  62.8   66.5
         2   Weight  19      0    99.5  100.026316  22.773933  50.5  84.0  99.5  112.5

            Max
         0   16
         1   72
         2  150
```

```
In [60]: import saspy
         import pandas as pd
         sas = saspy.SASsession(cfgname='winlocal')
         w_class = sas.sasdata("CLASS","SASHELP")
         code=sas.teach_me_SAS(1)
         w_class.columnInfo()
```

```
SAS Connection established. Subprocess id is 12800

proc contents data=SASHELP.CLASS ;ods select Variables;run;
```

```
In [61]: import saspy
         import pandas as pd
         sas = saspy.SASsession(cfgname='winlocal')
         %cd C:\SASCourse\Week14
         p_class = pd.read_sas('class.sas7bdat', format='sas7bdat', encoding="utf-8")
         p_class.describe()
```

```
SAS Connection established. Subprocess id is 7572

C:\SASCourse\Week14
SAS Connection terminated. Subprocess id was 11240
SAS Connection terminated. Subprocess id was 12908
```

Out[61]:
```
              Age     Height      Weight
count   19.000000  19.000000   19.000000
mean    13.315789  62.336842  100.026316
std      1.492672   5.127075   22.773933
min     11.000000  51.300000   50.500000
25%     12.000000  58.250000   84.250000
50%     13.000000  62.800000   99.500000
75%     14.500000  65.900000  112.250000
max     16.000000  72.000000  150.000000
```

In [62]:
```
import saspy
sas = saspy.SASsession(cfgname='winlocal')
```

```
SAS Connection established. Subprocess id is 12680
```

In [63]:
```
%%SAS
proc print data=sashelp.class (obs=5);
run;
```

Out[63]: <IPython.core.display.HTML object>

In [64]:
```
# List directories
import os, sys
path_x='/SASCourse'
dirs = os.listdir( path_x )
print(dirs)
```

```
['.git', '.gitignore', 'Assessments', 'Assignments', 'Clickable_URLs_SASTopics', 'have.csv', 'LI
```

In [65]:
```
import os
os.getcwd()
os.chdir('/SASCourse/Week1')
files = os.listdir(os.curdir)
print(files)
```

```
['Ex10_Data_step_view_etc.sas', 'Ex11_proc_print.sas', 'Ex12_Data_step_without_datalines.sas', '
```

In [66]:
```
import sys
sys.path
```

Out[66]:
```
['C:\\Users\\pmuhuri',
 'C:\\Users\\pmuhuri\\AppData\\Local\\Continuum\\anaconda3\\python37.zip',
 'C:\\Users\\pmuhuri\\AppData\\Local\\Continuum\\anaconda3\\DLLs',
 'C:\\Users\\pmuhuri\\AppData\\Local\\Continuum\\anaconda3\\lib',
 'C:\\Users\\pmuhuri\\AppData\\Local\\Continuum\\anaconda3',
```

```
 '',
 'C:\\Users\\pmuhuri\\AppData\\Roaming\\Python\\Python37\\site-packages',
 'C:\\Users\\pmuhuri\\AppData\\Local\\Continuum\\anaconda3\\lib\\site-packages',
 'C:\\Users\\pmuhuri\\AppData\\Local\\Continuum\\anaconda3\\lib\\site-packages\\win32',
 'C:\\Users\\pmuhuri\\AppData\\Local\\Continuum\\anaconda3\\lib\\site-packages\\win32\\
 'C:\\Users\\pmuhuri\\AppData\\Local\\Continuum\\anaconda3\\lib\\site-packages\\Pythonw
 'C:\\Users\\pmuhuri\\AppData\\Local\\Continuum\\anaconda3\\lib\\site-packages\\IPython
 'C:\\Users\\pmuhuri\\.ipython']
```

In [67]:
```
# https://thispointer.com/python-how-to-get-list-of-files-in-directory-and-sub-director
import os
def getListOfFiles(dirName):
    # create a list of file and sub directories
    # names in the given directory
    listOfFile = os.listdir(dirName)
    allFiles = list()
    # Iterate over all the entries
    for entry in listOfFile:
        # Create full path
        fullPath = os.path.join(dirName, entry)
        # If entry is a directory then get the list of files in this directory
        if os.path.isdir(fullPath):
            allFiles = allFiles + getListOfFiles(fullPath)
        else:
            allFiles.append(fullPath)

    return allFiles

import pandas as pd
dirName = '/SASCourse';
# Get the list of all files in directory tree at given path
listOfFiles = getListOfFiles(dirName)
df = pd.DataFrame(listOfFiles)
print(df)
```

```
                                                    0
0                        /SASCourse\.git\COMMIT_EDITMSG
1      /SASCourse\.git\Compiled_Macros\sasmacr.sas7bcat
2                                /SASCourse\.git\config
3                           /SASCourse\.git\description
4                            /SASCourse\.git\FETCH_HEAD
5                                  /SASCourse\.git\HEAD
6           /SASCourse\.git\hooks\applypatch-msg.sample
7               /SASCourse\.git\hooks\commit-msg.sample
8        /SASCourse\.git\hooks\fsmonitor-watchman.sample
9              /SASCourse\.git\hooks\post-update.sample
10          /SASCourse\.git\hooks\pre-applypatch.sample
11              /SASCourse\.git\hooks\pre-commit.sample
```

```
12                /SASCourse\.git\hooks\pre-push.sample
13              /SASCourse\.git\hooks\pre-rebase.sample
14             /SASCourse\.git\hooks\pre-receive.sample
15       /SASCourse\.git\hooks\prepare-commit-msg.sample
16                   /SASCourse\.git\hooks\update.sample
17                                /SASCourse\.git\index
18                           /SASCourse\.git\info\exclude
19                              /SASCourse\.git\logs\HEAD
20                /SASCourse\.git\logs\refs\heads\master
21          /SASCourse\.git\logs\refs\heads\SASCourse_rev
22          /SASCourse\.git\logs\refs\remotes\origin\HEAD
23        /SASCourse\.git\logs\refs\remotes\origin\master
24     /SASCourse\.git\logs\refs\remotes\origin\SASCo...
25     /SASCourse\.git\objects\00\a680e839a02561602f7...
26     /SASCourse\.git\objects\00\b3ae3a0a54838911c66...
27     /SASCourse\.git\objects\00\ea587a0c3f6071cbd26...
28     /SASCourse\.git\objects\00\f0a349fc2268af9b6df...
29     /SASCourse\.git\objects\01\0b96f8789c91acce5f3...
...                                                  ...
1131                /SASCourse\Week7\Week7_Handouts.docx
1132                 /SASCourse\Week7\Week7_Handouts.pdf
1133                /SASCourse\Week7\~$ek7_Handouts.docx
1134  /SASCourse\Week9\Ex10_create_macro_vars_call_s...
1135    /SASCourse\Week9\Ex11_create_macro_vars_sql.sas
1136    /SASCourse\Week9\Ex12_GetInfo_Host_Computer.sas
1137                      /SASCourse\Week9\Ex13_Macro.sas
1138       /SASCourse\Week9\Ex14_symlocal_symglobal.sas
1139             /SASCourse\Week9\Ex15_Indirect_ref.sas
1140      /SASCourse\Week9\Ex16_posi_key_para_macro.sas
1141                /SASCourse\Week9\Ex17_put_putlog.sas
1142     /SASCourse\Week9\Ex18_Week_9_List_of_Files.sas
1143       /SASCourse\Week9\Ex19_Macro_Vars_Resolve.sas
1144  /SASCourse\Week9\Ex1_Motivation_for_macro_vari...
1145                    /SASCourse\Week9\Ex20_SYMGET.sas
1146   /SASCourse\Week9\Ex21_Delete_Macro_Variables.sas
1147      /SASCourse\Week9\Ex22_Macro_vars_num_char.sas
1148                /SASCourse\Week9\Ex2_percent_let.sas
1149         /SASCourse\Week9\Ex3_Putlog_PercentPut.sas
1150         /SASCourse\Week9\Ex4_Text_Substitution.sas
1151                /SASCourse\Week9\Ex5_Lookup_mvar.sas
1152       /SASCourse\Week9\Ex6_Join_macro_var_text.sas
1153      /SASCourse\Week9\Ex7_indirect_reference_1.sas
1154               /SASCourse\Week9\Ex8_Global_Local.sas
1155       /SASCourse\Week9\Ex9_macro_vars_transfer.sas
1156           /SASCourse\Week9\Resolve_macro_vars.xlsx
1157                       /SASCourse\Week9\SomeData.txt
1158                /SASCourse\Week9\Week9_Handouts.docx
1159                 /SASCourse\Week9\Week9_Handouts.pdf
```

```
1160               /SASCourse\Week9\~$ek9_Handouts.docx

[1161 rows x 1 columns]
```

```
In [68]: import pandas as pd
         from pathlib import Path
         import time

         p = Path("/SASCourse/Week1")
         all_files = []
         for i in p.rglob('*.*'):
             all_files.append((i.name, i.parent, time.ctime(i.stat().st_ctime)))

         columns = ["File_Name", "Parent", "Created"]
         df = pd.DataFrame.from_records(all_files, columns=columns)

         print(df.iloc[:,0])
```

```
0              Ex10_Data_step_view_etc.sas
1                      Ex11_proc_print.sas
2     Ex12_Data_step_without_datalines.sas
3                   Ex13_Syntax_Errors.sas
4                Ex14_Pathname_Library.sas
5                             Ex1_DSPS.sas
6                         Ex2_Comments.sas
7                     Ex3_DataProcSteps.sas
8                          Ex4_DM_Clear.sas
9                       Ex5_proc_printto.sas
10           Ex6_Conpilation_Execution.sas
11       Ex7_Referencing_SAS_Data_Sets.sas
12                Ex8_Contents_all_ods.sas
13              Ex9_Contents_many_ways.sas
14                              PP_log.TXT
15                           PP_OUTPUT.TXT
16                     Week1_Handouts.docx
17                      Week1_Handouts.pdf
Name: File_Name, dtype: object
```

```
In [69]: from pathlib import Path
         dir =  Path('C:/Users/pmuhuri/Documents')
         files = dir.glob('*.docx')
         for i in files:
             print(i)
```

```
C:\Users\pmuhuri\Documents\Added_sspiauth_dll_to_PATH_environment_var.docx
C:\Users\pmuhuri\Documents\anacona_address.docx
C:\Users\pmuhuri\Documents\AnacondaPycharm.docx
```

```
C:\Users\pmuhuri\Documents\CCAS_Computer_Login_Issue.docx
C:\Users\pmuhuri\Documents\For_CCAS.docx
C:\Users\pmuhuri\Documents\In Dell Computer.docx
C:\Users\pmuhuri\Documents\JetBrain_AnacondaPycharm.docx
C:\Users\pmuhuri\Documents\Jupytor_SAS_Python_PyCharm.docx
C:\Users\pmuhuri\Documents\Licensed_SAS_JupyterLab.docx
C:\Users\pmuhuri\Documents\Pandas.docx
C:\Users\pmuhuri\Documents\Pop_issuues.docx
C:\Users\pmuhuri\Documents\Potental_Topics_Final_Exam_Spring_2019.docx
C:\Users\pmuhuri\Documents\Python as a calculator.docx
C:\Users\pmuhuri\Documents\py_workshops.docx
C:\Users\pmuhuri\Documents\SASpy.docx
C:\Users\pmuhuri\Documents\saspy_class.docx
C:\Users\pmuhuri\Documents\SAS_UE_power_off.docx
C:\Users\pmuhuri\Documents\Survival_Function.docx
C:\Users\pmuhuri\Documents\survival_models.docx
C:\Users\pmuhuri\Documents\Washington_Kali_Temple_Donation_Receipts_2018.docx
C:\Users\pmuhuri\Documents\Week11_Handouts_Fall2018.docx
C:\Users\pmuhuri\Documents\Working efficiently with JupyterLab Notebooks.docx
C:\Users\pmuhuri\Documents\~$ Dell Computer.docx
C:\Users\pmuhuri\Documents\~$SASpy.docx
C:\Users\pmuhuri\Documents\~$S_Topics_in_Papers_Blogs_11_08.docx
C:\Users\pmuhuri\Documents\~$_workshops.docx
```

In [ ]: