2.6_solutions

January 26, 2023

1 Exercises

1.0.1 reading and saving files

You will need to put the data folder `data_mixed` in the same directory as your notebook.

1. Import the os package and print a list of all the files in the data directory

[6]:
```
import os
os.listdir('./data')
```

[6]:
```
['CashFlow.xlsx',
 'DowJonesPrices - Copy.csv',
 'DowJonesPrices.csv',
 'fortune500.csv',
 'GermanDataClean_OUT.csv',
 'giltBondPrices.txt']
```

2. Open the file DowJonesPrices.csv using pandas and put the information in a dataframe.

[7]:
```
import pandas as pd
df=pd.read_csv('./data/DowJonesPrices.csv')
df.head()
```

[7]:

	Date	AAPL	AXP	BA	CAT	CSCO	\
0	2008-03-19	17.400475	36.908495	61.077398	59.969324	21.842601	
1	2008-03-20	17.883556	40.406014	62.199996	60.042524	22.110390	
2	2008-03-24	18.723591	41.662660	63.439005	61.840057	22.886976	
3	2008-03-25	18.918164	41.750537	63.114701	62.336210	22.985165	
4	2008-03-26	19.465665	39.861175	63.447322	62.921833	22.119317	

	CVX	DD	DIS	GE	…	NKE	PFE	\
0	63.884423	32.967519	28.272361	27.006786	…	27.788638	15.161948	
1	64.914187	32.953018	28.869664	28.448565	…	30.233569	15.147235	
2	65.538290	33.830312	28.996366	28.380270	…	31.047051	15.161948	
3	65.951755	34.294337	29.032567	28.281622	…	30.727951	15.287011	
4	66.279406	33.953570	28.742964	28.175386	…	29.784133	15.220801	

	PG	TRV	UNH	UTX	V	VZ	\

```
0  53.885155  38.364202  32.758480  57.891721  13.418429  22.318548
1  55.124812  39.504962  32.546541  58.102206  15.282759  22.947507
2  55.601600  39.422296  32.905916  59.221980  14.185536  23.487528
3  55.259901  39.480160  32.509681  59.070437  15.021516  23.443053
4  55.339365  39.306567  31.422339  58.607368  15.190136  22.960212

         WMT        XOM
0  42.863984  70.244690
1  44.932058  70.718922
2  45.269704  71.509308
3  44.780118  70.885317
4  44.653503  71.767228

[5 rows x 31 columns]
```

3. open the file CashFlow.xlsx and print a list with all the sheets in the file.

[9]:
```
my_xls = pd.ExcelFile('./data/CashFlow.xlsx')  # this opens the file for␣
↪inspection (all file, not just first sheet)
list_sheets=my_xls.sheet_names
list_sheets
```

[9]:
```
['CF_Scenario1', 'CF_Scenario2', 'CF_Scenario3']
```

4. Load the content of the first sheet of the file into a dataframe.

[10]:
```
# let's open the first
df_xls1=pd.read_excel('./data/CashFlow.xlsx', sheet_name='CF_Scenario1')
df_xls1
```

[10]:
```
    Year       CF  discount rate
0      0 -2000000           0.05
1      1   200000           0.05
2      2   200000           0.05
3      3   200000           0.05
4      4   200000           0.05
5      5   200000           0.05
6      6   200000           0.05
7      7   200000           0.05
8      8   200000           0.05
9      9   200000           0.05
10    10   200000           0.05
```

[ ]: