

Python for Financial Markets-Simple script to pull list of S&P 500 companies and save in a CSV file - Published by Ankush Garg

In this article I will cover how we can pull list of 500 S&P stocks from Wikipedia and save it in csv file.

Step 1: Get the wikipedia link for S&P 500 stocks

Step 2: use 'read_html' pandas built-in function to read the html page. If pandas is not present then install using the command pip install pandas

Step 3: Drop the column 'SEC filings' as this column contains link to SEC filing reports. When we write to csv these link will not function. That is why I am deleting this column

Step 4: Sort the list on Symbol (Ticker) in alphabetical order

Step 5: Save the pandas dataframe as csv file using the function to_csv

```
In [1]: import pandas as pd

# Read the data from html Link
wiki_data=pd.read_html('https://en.wikipedia.org/wiki/List_of_S%26P_500_companies')

# All data is stored in first cell
data = wiki_data[0]
display(data.head())

# drop columns 'SEC filings' Since we are dropping columns we have axis = 1 and inplace=True
data.drop(['SEC filings'], axis=1, inplace=True)

# Sort the dataframe on ticker in alphabetical ascending order
sorted_data = data.sort_values(by=['Symbol'], ascending=True)

# Convert the dataframe to csv file
# Index is False as we don't want to write index in csv file
sorted_data.to_csv('S&P500Tickers.csv', mode='w', index=False)
```

	Symbol	Security	SEC filings	GICS Sector	GICS Sub-Industry	Headquarters Location	Date first added	CIK	Founded
0	MMM	3M Company	reports	Industrials	Industrial Conglomerates	St. Paul, Minnesota	1976-08-09	66740	19
1	ABT	Abbott Laboratories	reports	Health Care	Health Care Equipment	North Chicago, Illinois	1964-03-31	1800	18
2	ABBV	AbbVie Inc.	reports	Health Care	Pharmaceuticals	North Chicago, Illinois	2012-12-31	1551152	20 (18)
3	ABMD	ABIOMED Inc	reports	Health Care	Health Care Equipment	Danvers, Massachusetts	2018-05-31	815094	19
4	ACN	Accenture plc	reports	Information Technology	IT Consulting & Other Services	Dublin, Ireland	2011-07-06	1467373	19