## Estimating Sectors – Readme

|  |  |  |
| --- | --- | --- |
| Library Name | Version used | Required |
| pandas | 0.16.2 | Yes |
| numpy | 1.11 | Yes |
| scikit-learn | 0.16 | Yes |
| wordcloud[[1]](#footnote-1) | 1.2 | Yes |
| Selenium[[2]](#footnote-2) | 3.0 | No (only used to fetch data) |
| yahoo-finance[[3]](#footnote-3) | 1.4 | No (only used to fetch data) |

The project is divided into 9 modules:

1. caps\_master.py: This is the main module and calls all other required modules. To run the model and generate requisite data and results, it is sufficient to run this file (takes around 10 minutes on a laptop).
2. caps\_common.py: This module contains functions which are used by other modules.
3. F1\_get\_fundamental\_data.py:.
4. F2\_get\_market\_data.py:
5. F3\_clean\_data.py:
6. F4\_create\_features.py:
7. F5\_create\_training\_data.py:
8. F6\_preprocess\_features.py:
9. F7\_train\_models.py:

1. https://github.com/amueller/word\_cloud [↑](#footnote-ref-1)
2. http://selenium-python.readthedocs.io/ [↑](#footnote-ref-2)
3. https://pypi.python.org/pypi/yahoo-finance [↑](#footnote-ref-3)