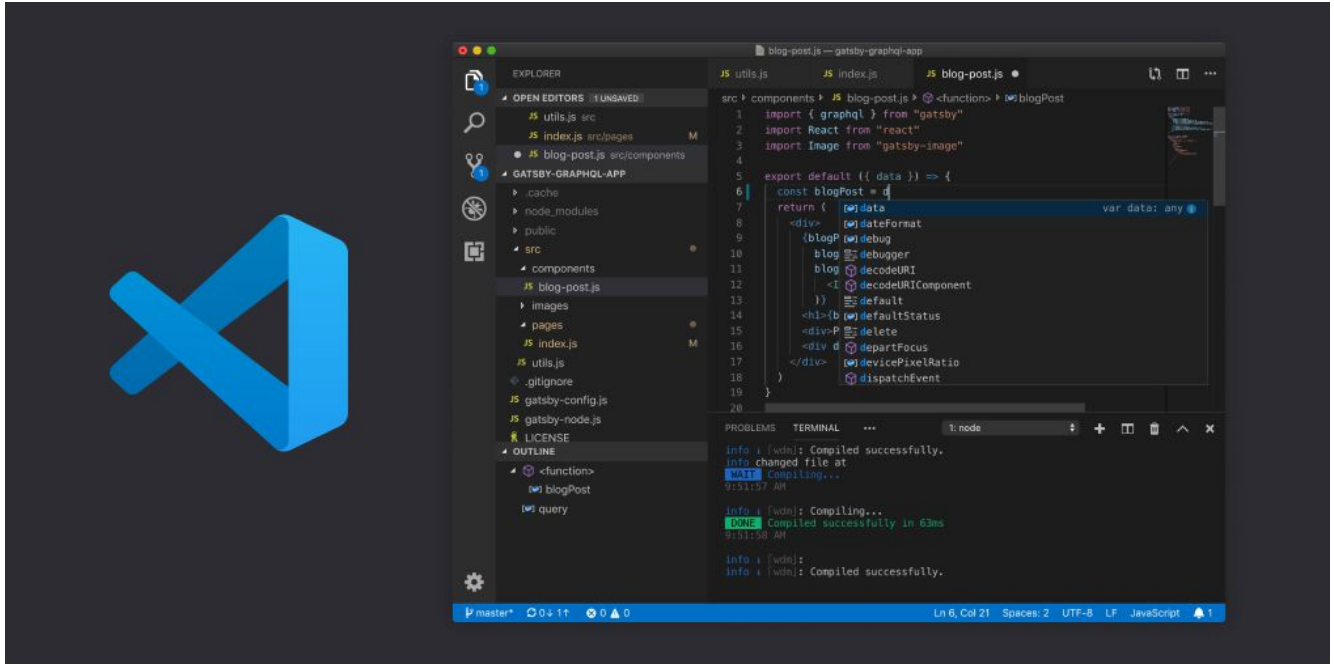
Anaconda

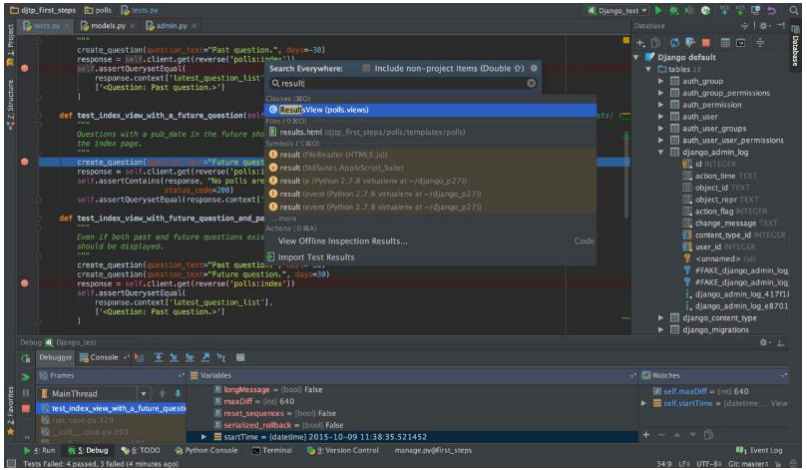
IDE

Visual Studio Code



https://code.visualstudio.com/download

PyCharm



https://www.jetbrains.com/ko-kr/pycharm/

Colab, Jupyter Notebook, Jupyter Lab

Pandas

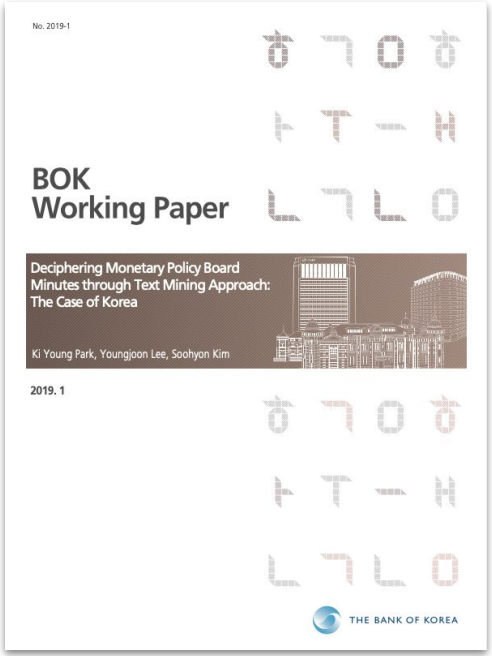
Numpy

Regular Expression

Deciphering Monetary Policy Board Minutes through Text Mining Approach: The Case of Korea

(텍스트 마이닝을 활용한 통화정책 분석)

논문



https://www.bok.or.kr/portal/bbs/P0002454/view.do?nttId=10049321&menuNo=200431&pageIndex=2

텍스트 수집

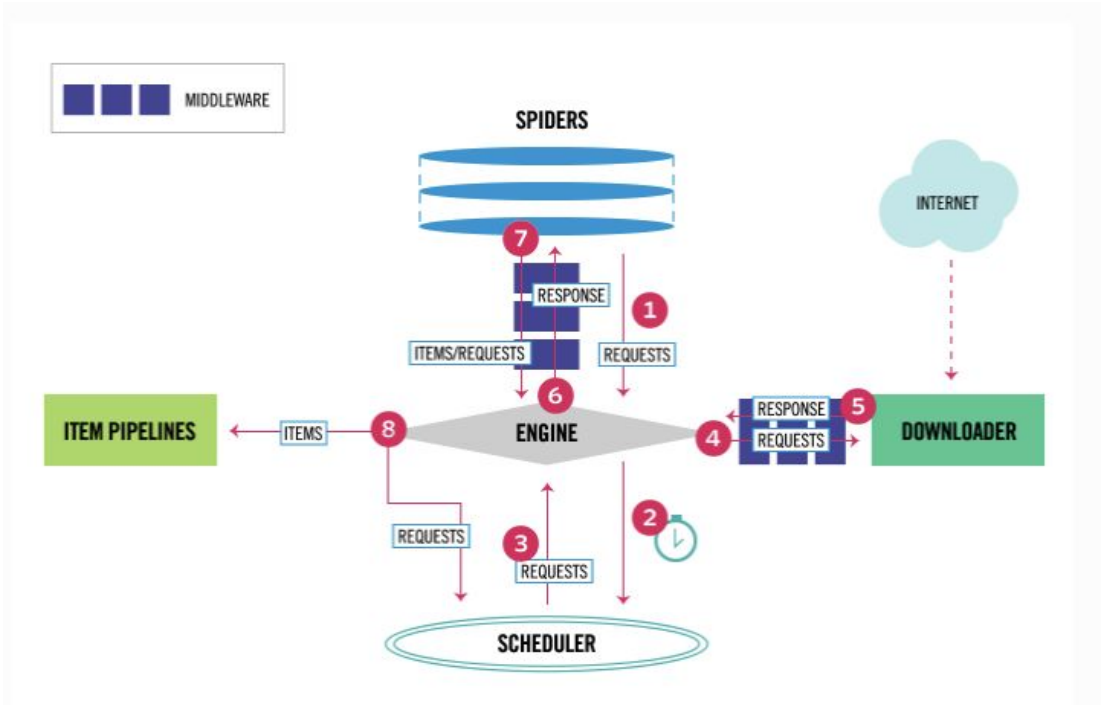
텍스트 수집

유효한 사전구축을 위해서 통화정책과 관련된 다량의 문서를 수집

- 금통위 의사록
- 금리 관련 뉴스
- 채권 애널리스트 분석 보고서
- 콜금리

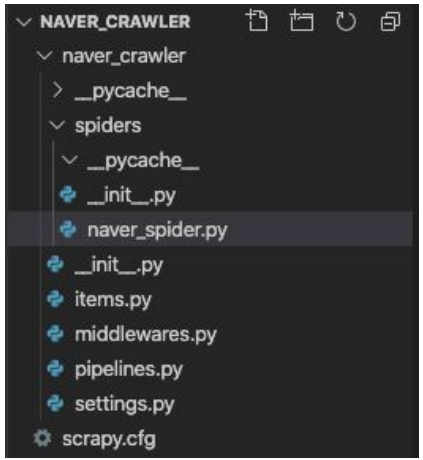
크롤링

Scrapy



텍스트 수집

- pip install scrapy
- scrapy startproject naver_crawler
- settings.py
- naver_spider.py under spiders




```python
import scrapy

class NaverSpider(scrapy.Spider):
    name = "naver"

    def start_requests(self):
        urls = [
            'https://news.naver.com/main/read.nhn?mode=LSD&mid=sec&sid1=102&oid=001&aid=0011729190'
        ]
        for url in urls:
            yield scrapy.Request(url=url, callback=self.parse)

    def parse(self, response):

        content = response.xpath("//div[@id='articleBodyContents']//text()").getall()
        self.log(content)
```