# 금융 데이터 관리사 강의자료

**이 문서의 링크:** [**bit.ly/3dHdqkl**](https://bit.ly/3dHdqkl)

# DAY 01 - 데이터 분석 환경과 데이터 수집

[DAY 01 메모](https://docs.google.com/document/d/1yDjV7xywLD2-T27dTaTbSrcUWkxuAz4hEMhkAJ_-p9k/edit?usp=sharing)

### 빅데이터

* [빅데이터 01 - 빅데이터와 데이터사이언스 - 슬라이드](https://drive.google.com/file/d/17ZAwpPfixEKiK3g1a2XO8PiLWWR4LtB5/view?usp=sharing)
* [빅데이터 02 - 빅데이터와 데이터사이언스 - 슬라이드](https://drive.google.com/file/d/1g_m1A2yb3G3zTQLPrPN7mus4ZFc1t4jT/view?usp=sharing)
* [빅데이터 03 - 기반기술과 데이터 수집 기술 - 슬라이드](https://drive.google.com/file/d/1PwJOAQxtqDgKCkK4josDVj-AUclNICxL/view?usp=sharing)
* [빅데이터 04 - 데이터의 종류와 기술 통계량 - 슬라이드](https://drive.google.com/file/d/1JsBTNEKOh0LeMuuYEl2uUAqSQ7OLW-to/view?usp=sharing)

### 파이썬 데이터 분석 환경과 도구

* [파이썬 소개 및 개발환경](https://drive.google.com/file/d/1UNLtkQw-v-pP_ykO-JDbHs0tF5WecIv4/view?usp=sharing)
* [Google Colab 튜토리얼](http://nbviewer.jupyter.org/4420df9e2c4660fd61828eb963ca9080)
* (참고) [로컬PC 개발환경 (윈도우 + 아나콘다)](https://nbviewer.jupyter.org/737719e50488821340b68a0d70bf0ffb)
* (참고) [파이썬 데이터 분석 환경 - 참고 사항들 (망분리)](https://drive.google.com/file/d/1P7PxBFLNlBaQTujm9Mra1PoxWsEFz5wl/view?usp=sharing) - 슬라이드

### 웹 크롤링 개요 (PDF 슬라이드)

* [웹 크롤링과 금융데이터 수집](https://drive.google.com/file/d/1zsC5dEG8S7cDjOMPUcYmXup3f4rXhhdD/view?usp=sharing)

### 오픈 API 개요

* [오픈 API 웹 기술 개요](https://drive.google.com/file/d/15LEymTE3QGSS9DjRcMs6nHzJ4ilKNO28/view?usp=sharing)

### 데이터 크롤링

* [데이터 크롤링 01 - 개요](https://nbviewer.jupyter.org/0bcce64a22feb3e2640e7086908fd812)
* [데이터 크롤링 02 - requests, BeautifulSoup](https://nbviewer.jupyter.org/f03474b4c1ea11b2a05af169f9db60d5)
* [데이터 크롤링 03 - read\_html() 활용](http://nbviewer.jupyter.org/a6c089a9f81d8fe37cadb2108453209a)

### **한국은행 OpenAPI**

* [한국은행 Open API - 개요](https://nbviewer.jupyter.org/dc04041d4afa792a0b7c7cfe0954a892)
* [한국은행 Open API - 100대 통계 지표](https://nbviewer.jupyter.org/54e0e507aa88075c97b8cc8fdaee6a46)
* [한국은행 Open API - 소비자 물가지수 생산자 물가지수](https://nbviewer.jupyter.org/4ab04bcea7e8fa2c2c39f61a07cffb1c)

### 금융감독원 - 금융상품 한눈에

* [금융상품 한눈에 01 - API 사용법](https://nbviewer.org/gist/FinanceData/14b3e81cbcdf9f29d32e4314b4da3aa3/notebook.ipynb)
* [금융상품 한눈에 02 - 전체 데이터 수집과 저장](http://nbviewer.jupyter.org/4f98465d84bbf210af83064208c4a843)

### 금융감독원 - 금융통계정보시스템

* [금융감독원 금융통계정보시스템 - OpenAPI 기본](http://nbviewer.jupyter.org/37e72180c8a3f1697e7d144a2a194f30) (API 사용하기)
* [금융감독원 금융통계정보시스템 - OpenAPI 고급](http://nbviewer.jupyter.org/37dd9106c8c932f872f4058868c205bf) (전체은행 BIS비율 데이터)

### 공공데이터

* [공공데이터포털 - 사업자등록정보 진위확인 및 상태조회 서비스](https://nbviewer.jupyter.org/69e07e0a140ea5b24a1878cc9c6302bd)

### 실습 Lab

* [Lab 파이썬 실습 - 구글 Colab](https://colab.research.google.com/drive/1D2oQ7-AZlS-XRz0206YzXJquaBLbL8ss?usp=sharing) -- [**실습**](https://colab.research.google.com/drive/14RnceaLyLenK4EN8-F2POOvWKiay14Vc?usp=sharing)
* [Lab 데이터 크롤링 - 네이버 증권 지수 데이터](https://colab.research.google.com/drive/1i4fUFt8EXIlMQQvdEiENMo96EFzYDPaE?usp=sharing) -- [**실습**](https://colab.research.google.com/drive/17o26ji9xgjeQLtJH42SymAuOJ1hrbalU?usp=sharing)
* [Lab 데이터 크롤링 - pd.read\_html()](https://colab.research.google.com/drive/1x2Is_Mij53812malocu2xT3hmwfKpWK4?usp=sharing) -- [**실습**](https://colab.research.google.com/drive/1w9LBvVAPoMRqTcKXfQzKAtfHluP84qpF?usp=sharing)
* [Lab 데이터 크롤링 - 개별종목 팩터 데이터 엑셀 저장](https://colab.research.google.com/drive/1griOkgoAIHgmzwWCsysTCNS3ZwVoAeKL?usp=sharing) -- [**실습**](https://colab.research.google.com/drive/1pCiSF0j3g0L0gxm27Qrr6s0fheWIH4UW?usp=sharing)
* [Lab 오픈API - 한국은행 OPEN API](https://colab.research.google.com/drive/1KlxPbTDwSLVkvp9MY5ptG1E14ziBCj8g?usp=sharing) -- [**실습**](https://colab.research.google.com/drive/1ueqt-iFGrzhTuEK48TFVovKwjwe6J5FO?usp=sharing)
* [Lab FRED 환율데이터 CSV](https://colab.research.google.com/drive/1bwZPyqJpVaCaGHYGcuvdHk2-OZzh3uli) -- 실습(구글 Colab -- 실습에 포함)
* [금감원 금융상품한눈에 오픈API.ipynb 실습](https://colab.research.google.com/drive/1-wgt_t0C9-NeoM-saI3p7SBIZZJL4oR6?usp=sharing)

# DAY 02 - 금융 데이터 시각화

[DAY 02 메모](https://docs.google.com/document/d/1F585wz5Z2rIzm3eVrS2RxpX4JrF7QY5pQ4rsDLkMiik/edit?usp=sharing)

### Tidy Data

* [Tidy Data (슬라이드 PDF)](https://drive.google.com/file/d/1My4cDxohmIR-bbHqSnYu2AzdRnhLo2YY/view?usp=sharing)
* [Tidy Data (노트북)](https://nbviewer.org/96be175a14b744dc1001129c2c7f98a0)
* '[Tidy Animated Verbs' in Python pandas](https://nbviewer.org/gist/FinanceData/8156bd53ea3a36c88cb146e43d4da411/notebook.ipynb)

### EDA 소개

* [EDA - 탐색적 데이터 분석](http://nbviewer.jupyter.org/908f06038af937ee1fddad2538e831fc)
* [인구조사 소득 예측 데이터셋 EDA](https://nbviewer.org/b23a2687eada48e6a5ff60d74e4cfb45)

### 데이터 시각화

* [데이터 시각화 - 데이터 시각화 - 차트 선택 방법](https://nbviewer.jupyter.org/1c7b228d495d23ba55b2092403b85d4b)
* [파이썬 데이터 시각화](https://drive.google.com/file/d/1-cSf9OtjjMVbXEG1suEQAp3GVMZzXyOK/view?usp=sharing) (슬라이드)

### matplotlib - 데이터 시각화

* [01 - Matplotlib](https://nbviewer.jupyter.org/3ae09e20d1431f5e417e84f89ae77e9b)
* [02 - Series and DataFrame](https://nbviewer.jupyter.org/a58e6453fbc8518b3f968fa2a6f301ce)
* [03 - 여러가지 그래프](https://nbviewer.jupyter.org/6b6c16134c5ff25c23300e03863550e7)
* [04 - 사례분석 100대 광고주](https://nbviewer.jupyter.org/793932e0aab322d062edb4b381c73aa8)
* [05 - mpld3 노트북](https://nbviewer.org/c49c405cf8525d5c15c6ef8012466681)
* [mplfinance 캔들스틱 차트](https://nbviewer.org/70cd14ad494ff4475cec709ebb132a18)

### seaborn - 데이터 시각화

* [데이터 시각화 - seaborn 01 앤스컴 콰르텟](https://nbviewer.org/8c7cbea387ba563f4ce04e2f39a30c47)
* [데이터 시각화 - seaborn 02 기본 표현](https://nbviewer.org/gist/FinanceData/5f623e4363fc741e092c694e70d567e4/notebook.ipynb)
* [데이터 시각화 - seaborn 03 다양한 형태의 데이터](https://nbviewer.org/gist/FinanceData/00f2f6da8cc6f4f08a6ab37d0ceb5f34/notebook.ipynb)
* [데이터 시각화 - seaborn 04 주가 데이터 시각화](https://nbviewer.org/gist/FinanceData/706d8c8581574593d09462bbafd9564e/notebook.ipynb)
* [**데이터 시각화 - seaborn 05 전체 살펴보기 (Seaborn Cheat Sheet)**](https://nbviewer.org/gist/FinanceData/379b56a5829fdb6da317d0ca3b438976/notebook.ipynb)

### plotly

* [plotly - express](https://nbviewer.org/995a636d7c556a830e52fcd8a80b627e)
* [plotly - pandas cufflink iplot](https://nbviewer.org/cb68361e77b99f4cf58fcf9b7e06b4c3)

### 실습 Lab

* [Lab 900-01 파이썬 데이터 시각화 - matplotlib.ipynb](https://colab.research.google.com/drive/135KZAVTQQCSaAl5lr6TTSFJ1e8BD2R4j?usp=sharing)
* [Lab 900-02 파이썬 데이터 시각화 - pandas.ipynb](https://colab.research.google.com/drive/1_kRpUaVijk_5GTAp4wyRNqOj9EDf0fDF?usp=sharing)
* [Lab 900-03 파이썬 데이터 시각화 - seaborn.ipynb](https://colab.research.google.com/drive/19c-ECmuyO7V83qTbov_LvBnkuzNHfq8Z?usp=sharing)
* [Lab 910-01 Seaborn 전체 살펴보기.ipynb](https://colab.research.google.com/drive/1vpmUm4UpqByncl0fAkTW-grB2XWekpCR?usp=sharing)
* [인구조사 소득 예측 데이터셋 EDA](https://nbviewer.org/b23a2687eada48e6a5ff60d74e4cfb45) - [**실습**](https://colab.research.google.com/drive/1gL4vxPtNStxuQOXv8Q3djXqtPNUoc5dD?usp=sharing)

# DAY 03 - 금융 데이터 전처리 및 데이터 분석

[DAY 03 메모](https://docs.google.com/document/d/1e_pnEfIm8GFk2ttKbujn5TyZkp6acIWO0mt-1ywgI5k/edit?usp=sharing)

### 파이썬 판다스 (pandas for finance)

* [pandas 데이터 기본](https://drive.google.com/file/d/1RRa-jhbZJaEXeiN7ZwU_iiXDlAR3lgZT/view?usp=sharing) (슬라이드, PDF)
* [pandas 데이터 분석](https://drive.google.com/file/d/1dv3kgq5vQbBy164zLSpJpTmcF_g05L-Y/view?usp=sharing) (슬라이드, PDF)

### Pandas Exercises, Practice

* [pandas 01 - Introduction](http://nbviewer.jupyter.org/88bc267cc41ee71a6edb9f81ac3d62d8)
* [pandas 02 - Series](http://nbviewer.jupyter.org/891e63d0ca58f3814116c68a011e8c1d)
* [pandas 03 - DataFrame](http://nbviewer.jupyter.org/402cc61a30da9eeba32f5e357ab2de83)
* [pandas 04 - DataFrame Indexing](http://nbviewer.jupyter.org/f5194e9dc142e194f40384c87175e1ff)
* [pandas 05 - DataFrame Timeseries](https://nbviewer.jupyter.org/89a1126acc5665840dc93ab8d9e4d14b)
* [pandas 06 - DataFrame merge, pivot](http://nbviewer.jupyter.org/5631bcad6b464c72c2afba824537b19d)

### 데이터 수집과 전처리

* [데이터 수집과 전처리](https://drive.google.com/file/d/1hkzstrdMGsoQ2X9gE76zR3o9XXzsfDLu/view?usp=sharing) (슬라이드)

### 데이터 전처리

* [데이터 전처리 - 개요](http://nbviewer.jupyter.org/e5520e0b8cc7b176042157ebac990189)
* [데이터 전처리 - 실습](http://nbviewer.jupyter.org/19ba530838ae5a13c520ffc675aec13b)
* [데이터 파일(CSV) 인코딩](https://nbviewer.org/1d1ec6d0655c450bdfc21dee9b225578) (참고)
* [자주 사용하는 문자열 전처리](https://nbviewer.jupyter.org/cd9521d89b9d17bb3124a5d18a5a114e) (참고)

### CreDB

* [CreDB 개요](http://nbviewer.org/285459a8e23b2810693b43e453dbb823)
* [CreDB 데이터셋 EDA](https://gist.github.com/FinanceData/e85151856c30c3229343198cf8d9109b)

### **실습 Lab**

* [데이터 전처리](http://nbviewer.jupyter.org/19ba530838ae5a13c520ffc675aec13b) **-** [**실습**](https://colab.research.google.com/drive/12cCsYSBfa1MAgtlYChFqCnAjPQ-JM8sE?usp=sharing)
* [캘리포니아 주택가격 데이터 - 이상치탐지](https://nbviewer.org/66a497a4785077a35accdc6ce0f301c7) **-** [**실습**](https://colab.research.google.com/drive/1zKPNH8GihuQipruSIUXqmt9x57_pnhe-?usp=sharing)
* [개인신용표본DB(light)](https://colab.research.google.com/drive/10CJ7_YW0v1P4M14LzycquR6ak6Lw6l5F?usp=sharing) - [**실습**](https://colab.research.google.com/drive/10amuhJKLXE-B3RYtYWZr0MrcbJAKV8x2?usp=sharing)

## 참고자료

#### [참고] 판다스 치트 시트

* [판다스 치트 시트](https://nbviewer.org/7f034a36406d68d6e2be92cc5ed4bec8) (노트북)
* <https://pandas.pydata.org/Pandas_Cheat_Sheet.pdf>

#### [참고] 자주 사용하는 예제 데이터셋

* [데이터셋 - 아이리스 (붓꽃)](http://nbviewer.jupyter.org/b7c2a073463ef59938f2eaaf9d2c169c)
* [데이터셋 - 당뇨병](https://nbviewer.org/6f3cc191a5df1fdeb3edf9d373df2e12)
* [데이터셋 - 보스턴 주택 가격](https://nbviewer.org/gist/FinanceData/87619bf3a4df4dd67fa37f5c9bf24605/boston-house-prices-dataset.ipynb)
* [데이터셋 - 캘리포니아 주택가격 데이터](https://nbviewer.org/66a497a4785077a35accdc6ce0f301c7)
* [데이터셋 - 타이타닉](http://nbviewer.jupyter.org/02b5aba7dbcf9f32bddd7c17be5d479a)
* [데이터셋 - MNIST](http://nbviewer.jupyter.org/340f44ff86afe3984ea53b02aa901e9b)
* [데이터셋 - 유방암 진단](https://nbviewer.org/3d40c508a4736726dd24b04a0293b7e7)
* [데이터셋 - 와인 품종](https://nbviewer.org/fd9564d7fa5f042e5f5c50d01b1589ee)

#### **[참고] 운영체제별 matplotlib + 한글**

* [matplotlib 한글 - Windows 아나콘다](https://nbviewer.org/344b3d4db72c92618e6cf477a84f4dcc)
* [matplotlib 한글 - Google Colab](https://nbviewer.org/3c17ba21151af8086bfe1850d147d14d)
* [matplotlib 한글 - MacOS](https://nbviewer.org/5059a7651fc35addccc27a5a58bf010e)
* [matplotlib 한글 - Linux](https://nbviewer.org/5d0534ea552350815abf8d4574271573)

## **Q&A 이승준**

* 한국금융연수원 겸임교수, FinanceData.KR 대표
* [plusjune@gmail.com](mailto:plusjune@gmail.com) , 010-9690-4558 (카카오톡)
* 페이스북 메신저 <https://m.me/plusjune>
* 페이스북 페이지 <https://fb.com/financedata>

#### **2022 FinanceData.KR** [**facebook.com/financedata**](http://facebook.com/financedata)