

데이터 수집 및 분석

이 부분이 데이터분석에서 가장 많은 시간이 필요한 부분입니다. 데이터를 수집해야하고 분석할 수 있는 모양으로 데이터 가공을 해야합니다. 다행히 일봉데이터는 데이터 크렌징은 필요없습니다. 하지만, 판다스 라이브러리를 이용하여 많은 가공이 필요합니다. 이 부분은 피쳐 엔지니어링이라고도 부릅니다.

일봉 데이터 수집은 웹크롤링으로 수집할 수 도 있고, 파이썬 라이브러리로도 수집이 가능합니다. 일봉 데이터에는 시종고저 값과 거래량 값이 기본적으로 제공됩니다. 우리는 이렇게 5개의 데이터를 이용하여 위 가설을 검정 해 볼 것 입니다.