

미국 금융시장과 경제정책지표

데이터 엔지니어링

2차 프로젝트 4팀 2조

이영호, 이보경, 장정빈, 황예원

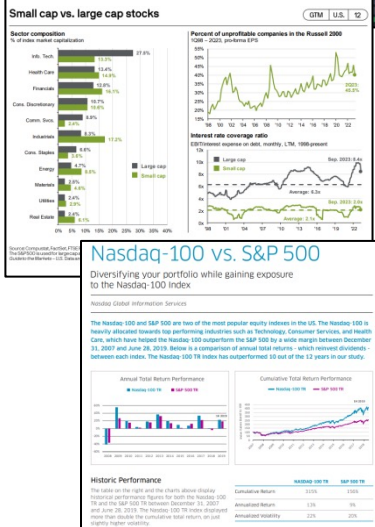
주제 선정 이유

1. 투자가 일상화되며 미국 주식 투자자수 증가 추세
2. 금리, 물가, 통화량 등 경제지표는 많지만 나의 투자에 어떤 영향을 끼치는지 파악하기 어려움
3. 미국 금융시장을 산업별로 분류하고 내가 투자한 종목이 어떤 경제지표와 가장 상관관계가 있는지 파악할 수 있는 대시보드를 구성

프로젝트 목표



1. 현재 많은 사람들이 사용 중인 금융 데이터 시각화 서비스와 비슷한
혹은 경우에 따라서는 더 유효한 대시보드



2. 이후에 다른 경제관련 데이터를 추가하기 쉽고
다른 서비스로 확장할 수 있는 대시보드

스키마 & 테이블

RAW_DATA

detail	
Company	VARCHAR(50)
Ticker	VARCHAR(50)
GICS sector	VARCHAR(50)
GICS Sub-Industry	VARCHAR(100)
Market Cap	BIGINT

prices	
Ticker	VARCHAR(50)
date	DATE
price	FLOAT

detail_sp	
Ticker	VARCHAR(50)
Company	VARCHAR(50)
GICS sector	VARCHAR(50)
GICS Sub-Industry	VARCHAR(100)
Headquarters Location	VARCHAR(100)
Date added	DATE
CIK	BIGINT
Market Cap	BIGINT

index	
date	DATE
CPI	FLOAT
M2	FLOAT
FEDFUNDS	FLOAT
HIGHYIELD	FLOAT

nasdaq_and_sp	
Date added	DATE
NASDAQ	FLOAT
SP500	FLOAT

RAW_DATA 스키마 아래

Prices: 미국 상장종목들의 20년치 일별 주가 데이터

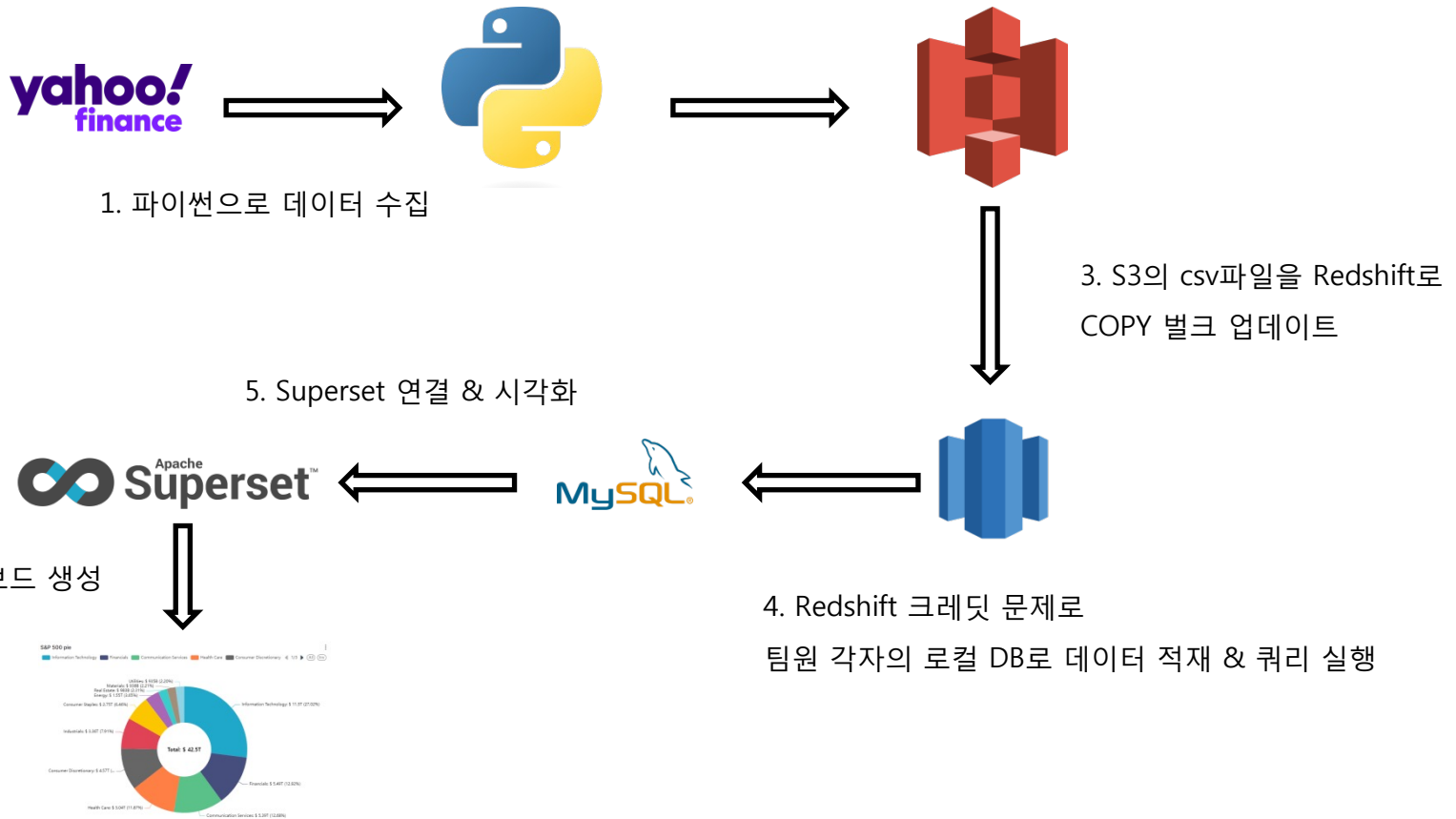
Index: 미국 경제 보조 지표

Nasdaq_and_SP: 미국 나스닥과 S&P500 지수

Detail, Detail_sp: 나스닥과 S&P500 상장 종목들에 대한 상세정보

활용 기술

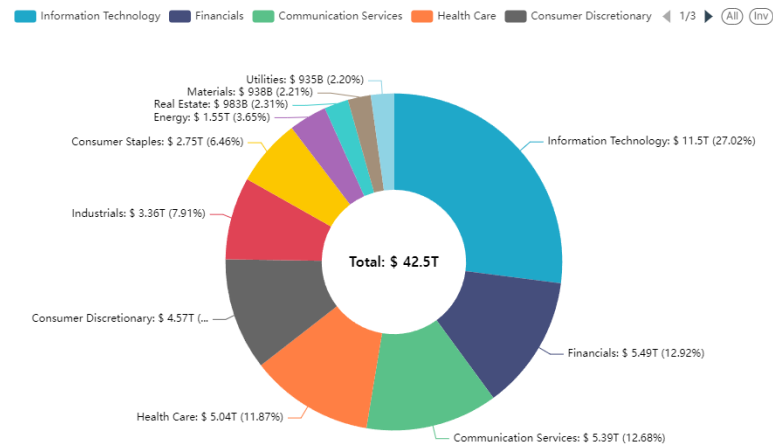
2. 파이썬(판다스)으로 전처리 후
S3에 csv 파일로 적재



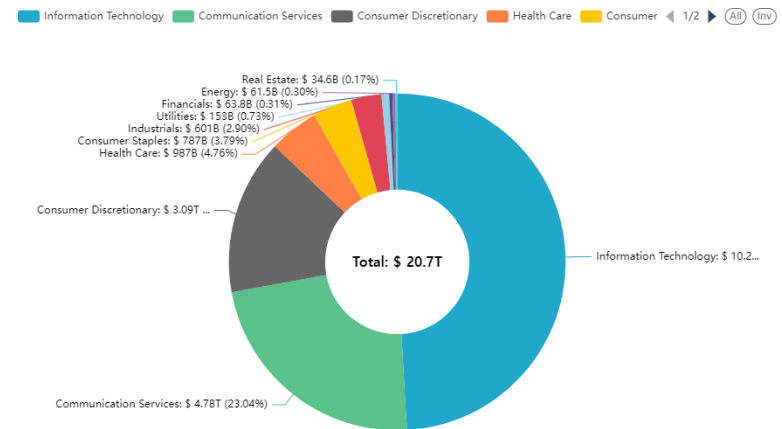
결과물

NASDAQ 100 and S&P 500

S&P 500 pie

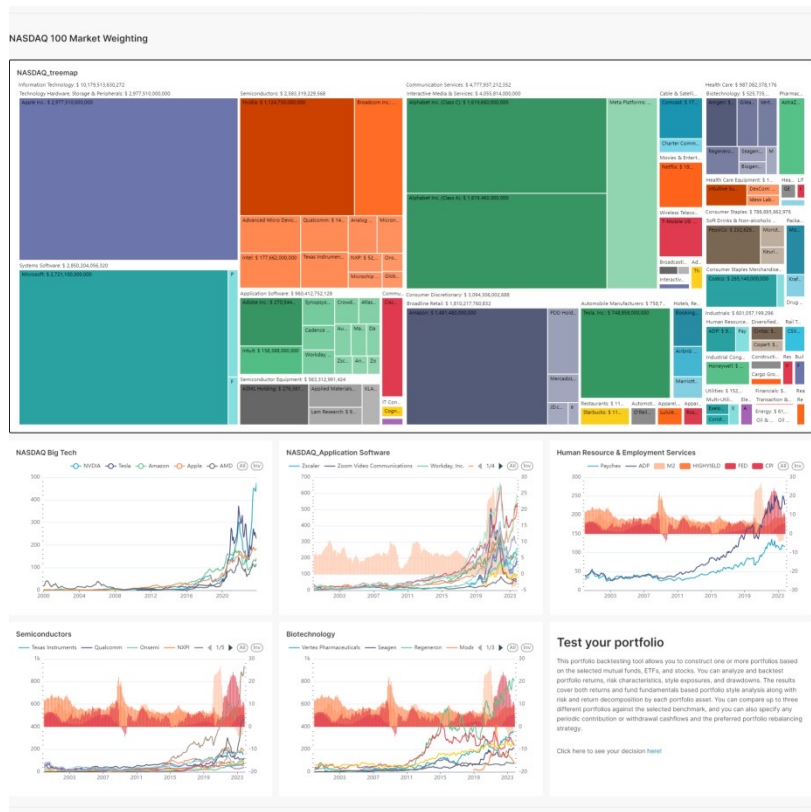


NASDAQ 100 pie



나스닥과 S&P500의 산업별 시가총액 비중

결과물



미국 금융시장 산업별 대분류, 소분류, 기업별 Treemap 차트

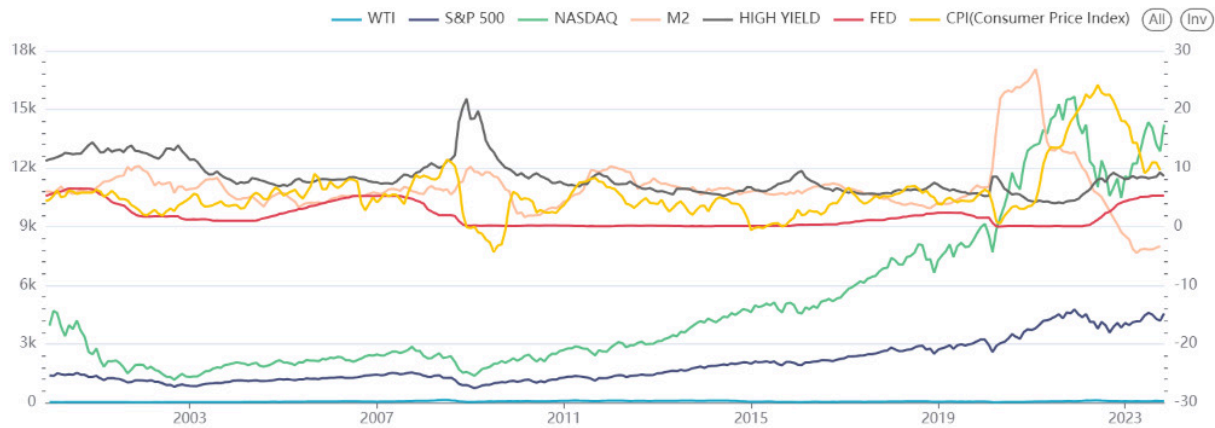
내가 투자한 종목들과 경제지표들과의 상관관계

나의 포트폴리오를 백테스팅할 수 있는 기능 첨부

결과물

Decision Making

Investment Decision Metrics



국제유가, 통화량, 연준 금리, 물가, 하이일드 등의 보조 지표

기대되는 효과

1. 금융시장이 어떤 산업대분류 소분류로 구분되고 각 산업별 비중이 어떻게 되는지 어떤 기업들로 구성되어있는지 한 눈에 살펴볼 수 있습니다.
2. 각 기업별 주가 추이를 Dual Y-axis를 이용하여 금리나 통화량 등과의 상관관계를 한눈에 살필 수 있습니다.
3. 다른 보조지표를 손쉽게 추가·제거해 새로운 차트를 생성할 수 있습니다.
4. 시계열 경제 지표 데이터를 수집할 수만 있다면 연도, 월, 일별로 전처리해 기존의 Dataset에 merge하거나 JOIN하여 시각화할 수 있습니다.

개선점

1. 금융 데이터들이 파편화되어 있고 시계열이 제각각이라 전처리 작업이 매우 힘들었습니다. 때문에 적시성이 매우 중요함에도 가장 필수적인 자동화를 구현하지 못했습니다.
2. 나스닥 100과 S&P 500은 비교적 기업의 수가 적지만 기타 모든 기업들을 대상으로 대시보드를 구성하면 Superset 대시보드 진입 시 쿼리문을 실행하기 때문에 속도가 매우 느립니다. 테이블 디자인을 조금 더 고민해봐야 할 것 같습니다.

◀ 감사합니다. ▶