

Brainstorm

- Valoración de empresas de NASDAQ o cualquier otro mercado de cotización
- Comparación de la valoración de estas empresas con sus valoraciones en bolsa
- Buscamos un dataset con un seguimiento de las empresas donde se vea que hay una diferencia entre la valoración que hemos hecho y su cotización en bolsa
- Se buscarán empresas que estimemos que están infravaloradas en bolsa en al menos un 20% para descubrir oportunidades de inversión. Y se hará una predicción de crecimiento
- Se puede filtrar por tipo de industria

Diseño del DAaaS

Mi estrategia es ofrecer un servicio de inversión en empresas .Se podría hacer una web para usuarios particulares pero también ofrecer un API al que empresas especializadas se puedan conectar

Necesitaremos:

- Histórico de datos de empresas descargado de la API de la SEC(<https://www.sec.gov/edgar/sec-api-documentation>)
- Noticias que puedan ser relevantes sobre el tipo de industria o las empresas que finalmente sean seleccionadas

Arquitectura DAaaS

- Componente para descargar los datos de la API (CSV) y llevarlos a mi google cloud
- Conexión a la API de la SEC(<https://www.sec.gov/edgar/sec-api-documentation>)
- Base de datos PostgreSQL para almacenar los datos recogidos
- Hadoop para procesar esos datos
- Componente que me haga la valoración y la comparación de esos datos y me de el resultado de las empresas que se ajusten a mi requerimiento
- Crawler que me busque noticias relevantes sobre las empresas elegidas y sobre los sectores a los que pertenecen las empresas
- Componente que con las noticias que se hayan encontrado finalice el proceso de valoración de la empresa
- Base de datos PostgreSQL que me guarde los resultados finales
- Servidor donde alojar mi web y mi API

DAaaS Operating Model Design and Rollout

1. Todos los días al cierre de mercado el componente se descarga los datos de la API y manda el CSV resultante al bucket de Google Cloud
2. Todos los días a las 8:00 ejecutamos manualmente el dataproc y el resultado lo almacenamos en la base de datos PostgreSQL
3. Ejecutamos el componente que nos hace las valoraciones
4. Ejecutamos el crawler buscando noticias relevantes
5. Ejecutamos el segundo algoritmo que nos da el resultado final y nos llevamos los resultados a la otra base de datos SQL
6. Ejecutaremos las queries requeridas en la base de datos
7. Nuestra web y nuestra API ya puede hacer llamadas a nuestra base de datos
8. También podremos usar la base de datos resultante para hacer presentaciones internas de nuestra empresa

Diagrama:

