# CARLES MATOSES MIRET

Diseño del DAaaS

### Definición la estrategia del DAaaS

El objetivo que se quiere conseguir con el trabajo es obtener predicciones de distintos valores en bolsa para comprar y vender acciones, tratando de obtener el máximo beneficio.

Las principales fuentes de origen van a ser las siguientes:

* Alpha vantage: <https://www.alphavantage.co/documentation/> . A través de su API obtendremos:
  + Valor de las acciones: apertura, cierre, máximo, mínimo y volumen
  + Indicadores técnicos: RSI, MACD, Volumen, etc.
  + Datos macroeconómicos: inflación, tipos de interés, datos de empleo de los Estados Unidos, datos de nóminas no agrícolas, etc
  + Estos datos se pueden obtener a nivel intradía, diarios, semanales y mensuales
  + Existe más de 20 años de histórico de datos de las principales acciones y divisas
* Yahoo finance: <https://pypi.org/project/yfinance/> . A través de su API obtendremos:
  + Cuenta de resultados por trimestre y las expectativas que tenían los analistas
  + Balance de situación de la empresa
  + Recomendaciones de los analistas (fuerte compra, compra, neutral, venta, fuerte venta)
  + Precio objetivo de las acciones
  + Máximos accionistas de las empresas
  + En general, todos los datos fundamentales de las empresas y datos sobre distintos analistas e inversores

Estos datos los obtendremos con scripts de Python, los almacenaremos en formato json, transformaremos en formato tabular los datos que necesitemos y vayamos a trabajar, para finalmente utilizar modelos de machine learning que realicen el pronóstico del futuro valor de las acciones y poder analizarlo en Power BI para decidir en qué empresas a invertir. A futuro, se trataría de automatizar por API la compra y venta de las acciones.

### Arquitectura DAaaS

Principalmente utilizaremos herramientas Cloud para poder hacer el proyecto escalable y para que al mismo tiempo sea un entorno de colaboración con otros compañeros.

* Utilizaremos **CLOUD FUNCTIONS** de Google Cloud para que se ejecuten scripts de Python que obtengan los datos de las dos APIs nombradas anteriormente
* Los datos obtenidos se almacenarán en formaton json en **CLOUD STORAGE** para tener siempre el máximo número de datos posible y luego extraer los necesarios.
* A través de **DATAPROC** transformaremos los json y los almacenaremos en formato tabular en **BIGQUERY**. En esta transformación crearemos distintas tablas, que entre ella se podrán relacionar y almacenar en el datawarehouse.
* Con los datos almacenados, utilizaremos las funcionalidades que nos ofrece **VERTEX AI** para realizar los modelos predictivos. Luego devolveremos los datos a Bigquery para almacenar los resultados y enriquecer la información que ya teníamos almacenada.
* Finalmente, a través de una **API GATEWAY** podremos conectar nuestro **POWER BI** on premise con Bigquery. Con Power BI realizaremos los análisis necesarios para la toma de decisiones en la compra y venta de acciones.
* Todos estos procesos los realizaremos con un orquestador pendiente de valorar pros y contras: **AIRFLOW** o **CLOUD COMPOSER** (es airflow dentro del entorno de Google Cloud)

### DAaaS Operating Model Design and Rollout

* La extracción de los datos técnicos se realizará todos los días
* Habrá un trigger que cada vez que haya algún cambio de recomendación de los analistas o de los datos fundamentales de la empresa, se los traerá.
* Tendremos un calendario que se irá actualizando para saber cuando presenta resultados cada empresa. Igualmente, también tendremos un trigger para cuando tengamos nuevos datos (los resultados) nos los traigamos.

### Desarrollo de la plataforma DAaaS. (ligera descripción del desarrollo)

Link a Diagrama:

