# Objetivos

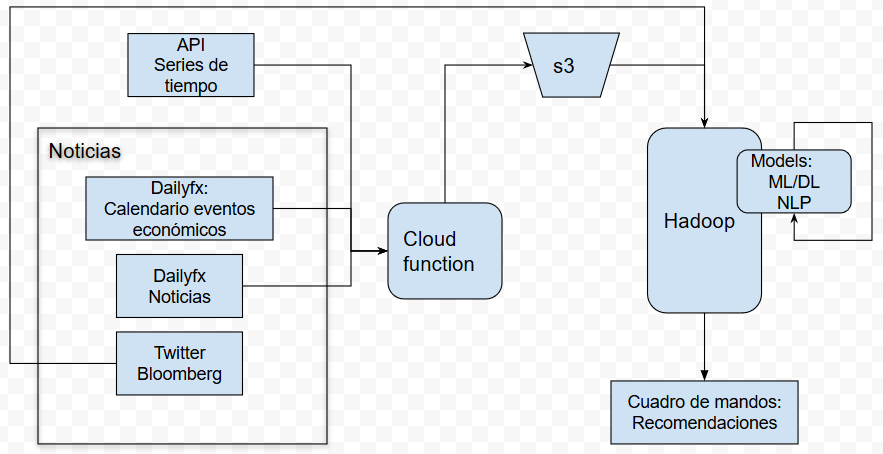
* Desarrollar un flujo que tome en cuenta la ingesta y procesamiento de series de tiempo así como data sets adicionales para enriquecimiento del análisis.
* Nos enfocaremos en el mercado de divisas por la disponibilidad y variedad de información que incluye series de tiempo y noticias, lo que nos permitirá aplicar técnicas de scrapping, machine learning, deep learning y nlp.
* Obtener una plataforma que permita desarrollar negocios dentro del mercado de divisas (sugerencias de trading basadas en AI, análisis, etc.) así como en el futuro poder enfocarnos a análisis de otros datos relacionados con series de tiempo como son los sensores.

# Definir Data Sets

1. Los datos los obtendremos de la API Forexconnect. Pertenece a una empresa que se dedica a la intermediación y a través de esta API permite que programadores desarrollen servicios de trading de divisas.
2. Haremos scrapping de la página de noticias y calendario de eventos económicos de [www.dailyfx.com](http://www.dailyfx.com) que publica información sobre el mercado financiero para captar clientes.
3. Tweets de páginas de noticias financieras como [Bloomberg](https://twitter.com/business).

# Arquitectura, almacenamiento e ingesta

La información se almacenará en Hadoop en un proveedor de la nube entre gcloud, aws o azure y procesada con Spark y Python.



# ETL y cálculo de métricas con Spark

Analizaremos los datos y haremos predicción con modelos ARIMA y redes neuronales basados en series de tiempo y nlp.

# Visualización de las métricas

El dashboard incluirá sugerencias de operaciones de trading basadas en los diferentes modelos e indicando probabilidad de éxito.