# BOOTCAMP DATA SCIENCE

MEMORIA PROYECTO FINAL: AGRO-TECH ***CropIT***

# GRUPO: Tanya Moscoso, Germán Fernández

## DESCRIPCIÓN DEL PROYECTO

CropIT es una Agro-TECH especializada en el desarrollo de soluciones Big Data - ámbito de la Agricultura de Precisión

Enfocada en la toma de decisiones basada en el manejo y tratamiento de los datos permite maximizar el rendimiento de los cultivos, a la vez que permite la optimización de los recursos empleados en la producción (agua, fertilizantes, pesticidas, etc.)

Destinada a:

* Pequeños y medianos agricultores que buscan mejorar las condiciones del sistema productivo a la vez que profundizar en el conocimiento de sus explotaciones.
* Técnicos de campo que realizan el manejo de una o varias fincas/explotaciones agrícolas.

## VISIÓN Y MISIÓN

VISIÓN:

* Problema: desarrollo anormal de los cultivos.   
  Solución: llevar a cabo un análisis de suelo para detectar posibles causas.   
    
  **¿Qué acciones llevar a cabo?** Necesidad de una interpretación visual y simplificada de los resultados de dicho análisis 🡪 dashboard de analítica.   
  **¿Estrategia a implementar?** Va en función del caso de uso (soluciones tailor-made).   
    
  **¿Qué puedes encontrar actualmente en el mercado?** Herramientas o literatura con ‘blanket recommendations’, etc. Heterogeneidad de fuentes.   
  **¿Qué podemos ofrecerte?** Planificación de las aportaciones de nutrientes/fertilizantes. Diseño del calendario óptimo de riesgo adaptado a las necesidades de tu cultivo/s. Eficiencia energética de la explotación – autoabastecimiento mediante fotovoltaica.   
    
  **¿Qué nos diferencia?** Combinamos la experiencia y el know-how del usuario respecto al manejo agrícola, con la analítica de sus propios datos (análisis de laboratorio, sensores en cultivos – IoT,…) para elaborar soluciones enfocadas a maximizar el rendimiento y optimizar los recursos. Evolución hacia agronomía data-driven.

MISIÓN:

* Convertirnos en referentes del sector, ampliar alcance y posibles aplicaciones, dar solución a los diferentes retos planteados por el cliente.

## ORIGEN DE LOS DATOS

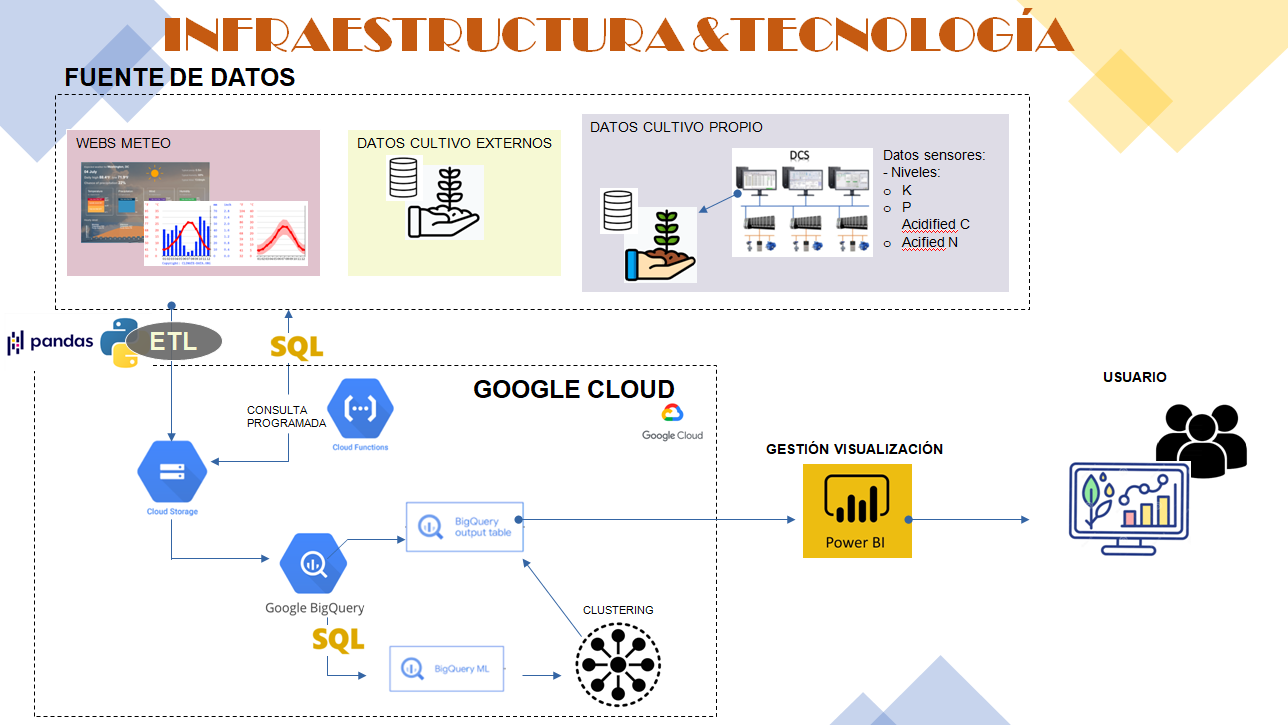
Para la toma de datos nos hemos basado en 3 fuentes principales:

1. Base de datos de propiedades de suelos 🡪 AfSIS DB. Fuente pública “open data”. Valores reales.
2. Datos climatología. Parte de valores reales y parte generados aleatoriamente para hacer coincidir con los datos de suelos utilizados.
3. Características paneles fotovoltaicos. Datos de proveedores de paneles fotovoltaicos.

## INFRAESTRUCTURA Y TECNOLOGÍA

La infraestructura ha sido dividida en 3 bloques generales:

1. Fuente de datos. Especificada en el apartado 3.
2. Ecosistema de Google Cloud para la gestión de los datos.
   1. Cloud Storage.
      1. Almacenamiento de datos.
   2. Cloud Functions.
      1. Consultas programadas para la actualización periódica de la base de datos.
   3. Google BigQuery.
      1. Creación de tablas de salida para usar en PowerBI.
      2. Generación de modelos de ML para clusterizar los datos de suelo.
3. Gestión de visualización con PowerBI.



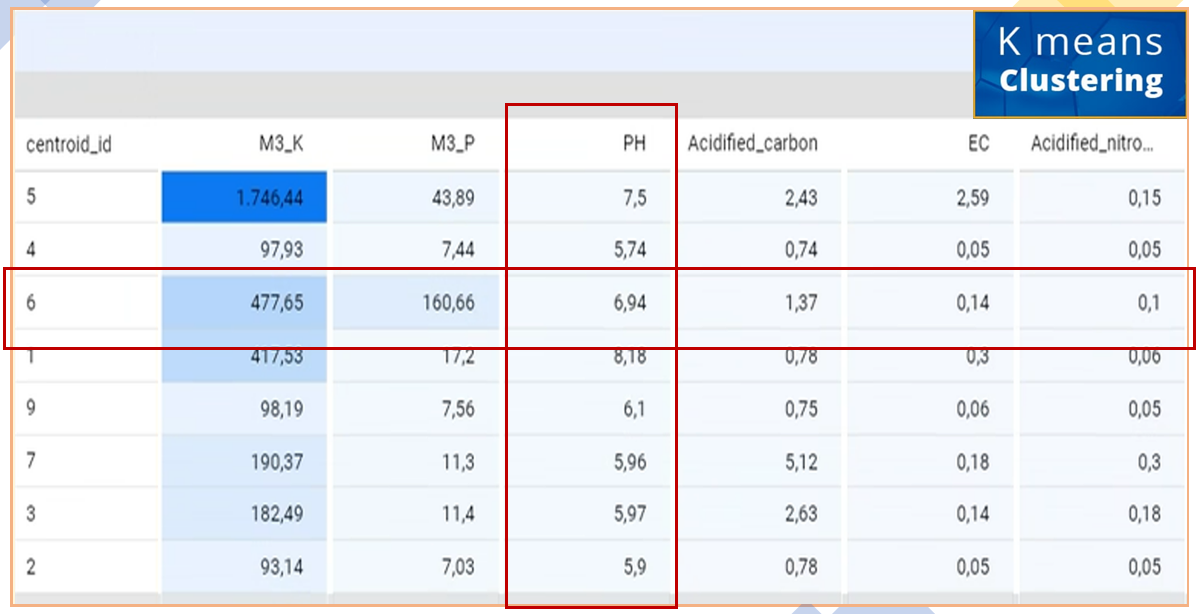
## DEMO

Se adjunta archivo PowerBI con la presentación de la memoria.

## CLUSTERING

Hemos usado un modelo de clustering para intentar definir los niveles de nutrientes óptimos para conseguir un pH neutro, o lo más cercano a él. Vemos que con un modelo k-means para 9 clusters, el cluster nº6 tiene un pH muy próximo a 7, nuestro valor objetivo.

Teniendo en cuenta estos parámetros podemos monitorizar nuestro cultivo para asegurar que todos los valores de nutrientes permanecen en los rangos requeridos y, en caso de necesitarlo, aportando nutrientes si detectamos carencia en alguno de los valores.



## PREVISIÓN FUTURA

A continuación se analizan varios puntos posibles de crecimiento futuro:

DESARROLLO Y PROSPECCIONES A LARGO PLAZO

* Gestión integral del análisis: asesorar en la toma correcta del dato (ofrecer el diseño e instalación de la red de sensores mediante IoT), y su posterior tratamiento y análisis hasta la consecución del modelo de manejo más óptimo.
* Diseño de futuras explotaciones. Si eres nuevo en el sector, nuestras herramientas te permitirán realizar un análisis detallado con el objetivo de elegir correctamente tu plantación, rotaciones, fechas de siembra/recolección, y el manejo integral de la misma.
* En caso de falta de recursos para acometer la mejora a nivel de campo y de sistemas, gestionar con nosotros la solicitud de ayudas/ financiación que mejor se adapte a cada situación particular.

DATA

* Posibilidad de crecimiento de la base de datos: trabajar no sólo con open datasets, si no retroalimentar los modelos con datos experimentales y en tiempo real de distintas parcelas, y los datos que pueda aportar el propio agricultor, para enriquecer y robustecer la base de datos.