

Data Mining para Agricultura

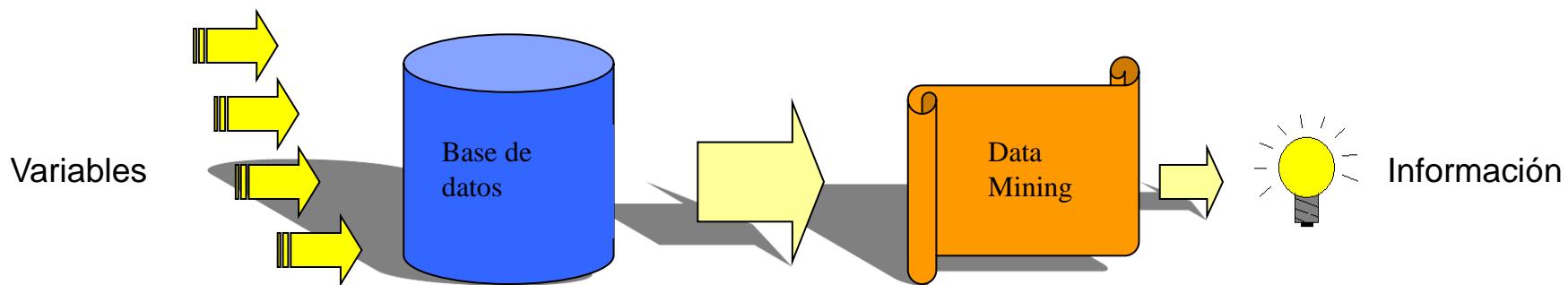
Marcelo Ferreyra





¿Qué es Data Mining?

Es el conjunto de herramientas y métodos que se usa para obtener información a partir de una gran cantidad de datos

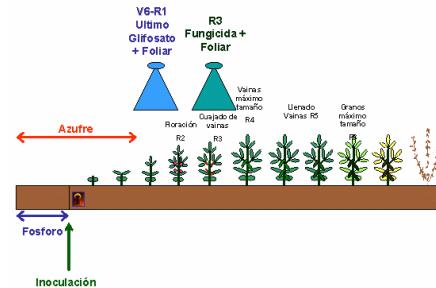




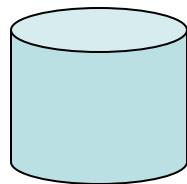
Los tres pilares del DM



Analistas



Conocedores del negocio



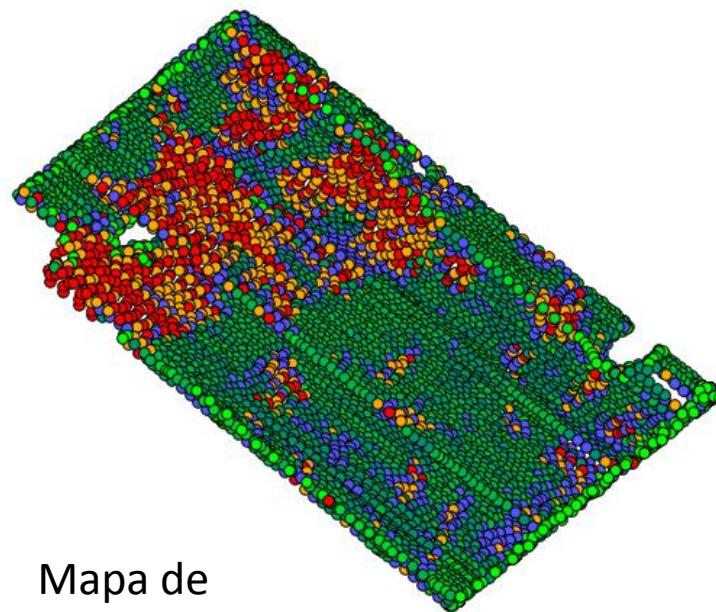
Datos



Objetivo del DM en AP

Se utiliza Data Mining en Agricultura para explicar la variabilidad

¿Qué variables condicionan el rendimiento en las distintas zonas y de qué manera lo hacen?





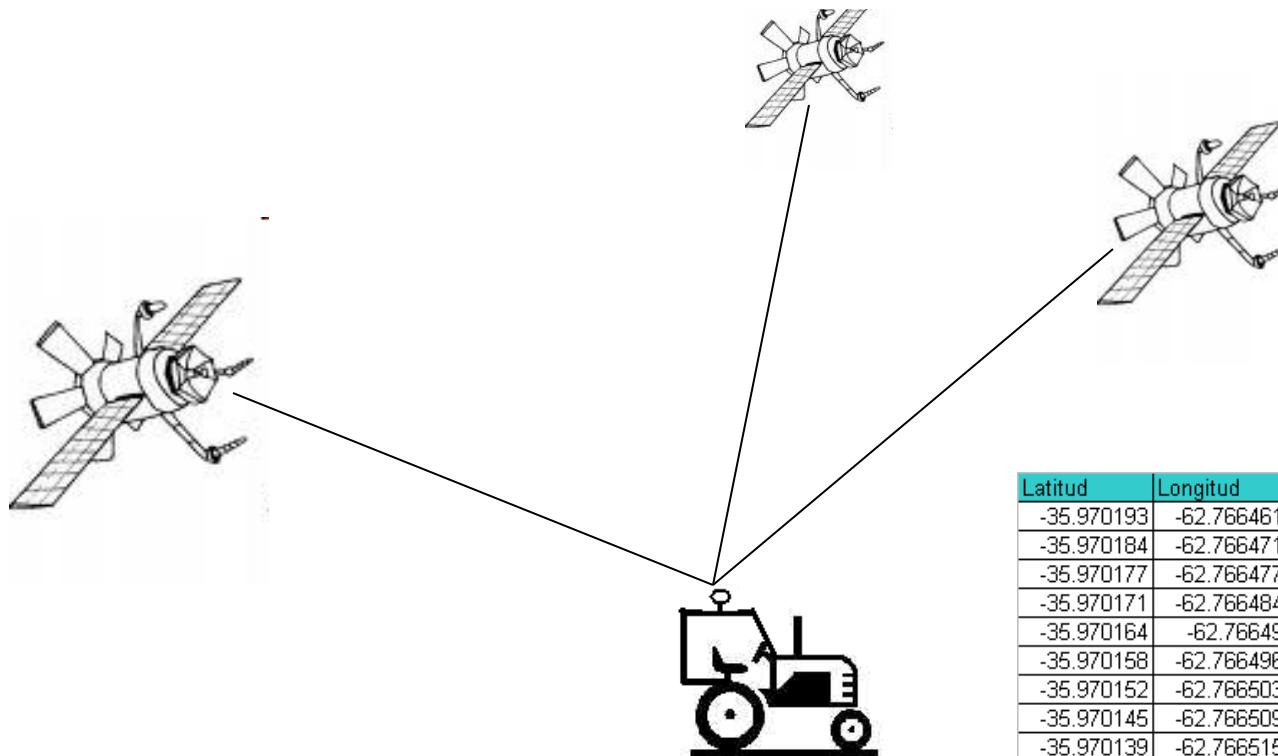
Datos en AP

La agricultura de precisión se nutre de datos de varias fuentes:

- Clima. Precipitaciones, temperaturas, vientos, radiación
- Propiedades físicas y químicas del suelo: nutrientes, Ph, salinidad, textura, humedad, % de arena
- Condiciones del cultivo: variedad, densidad y profundidad de siembra
- Aplicación de insumos: fertilizantes, herbicidas, insecticidas
- Rendimiento



Datos Georeferenciados





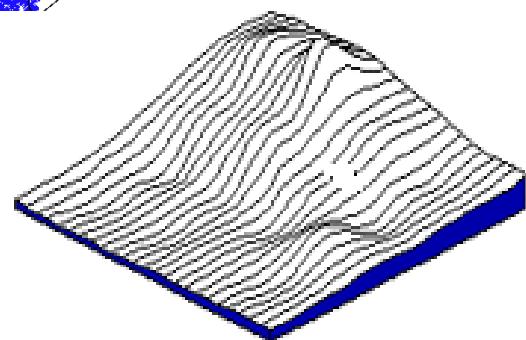
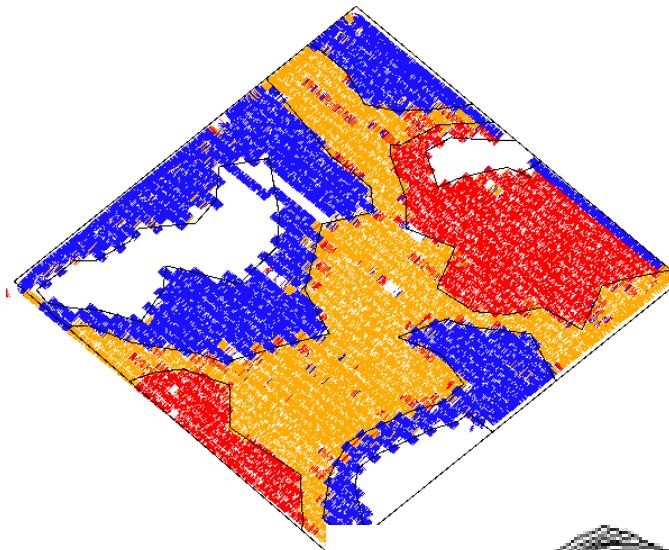
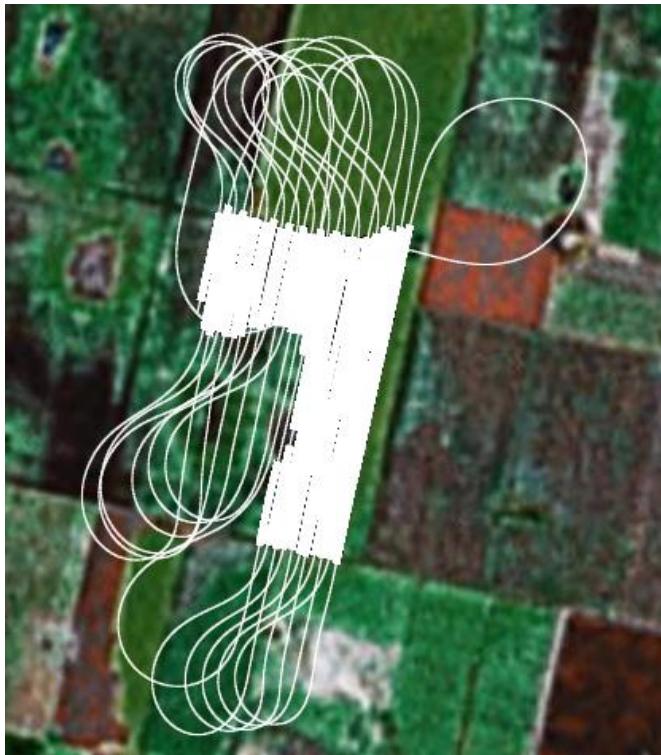
Calidad de los Datos

A fin de obtener datos de alta calidad es necesario controlar los errores. Las fuentes de errores más comunes son:

- Problemas mecánicos
- Problemas electrónicos
- Equivocaciones humanas



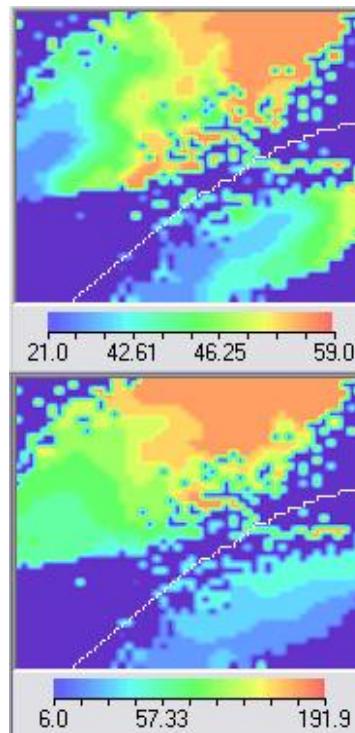
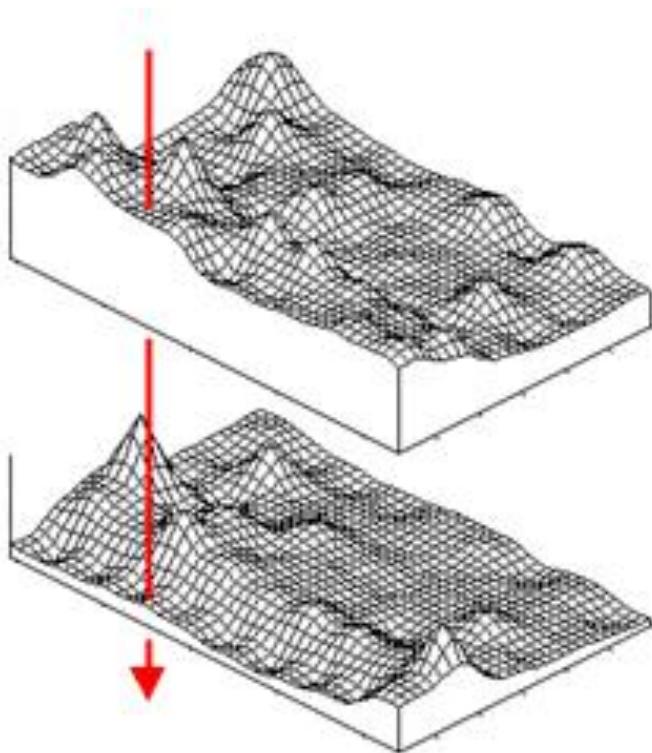
Visualizando los Datos





Comparando Mapas

¿Qué tan parecidos son estos mapas?

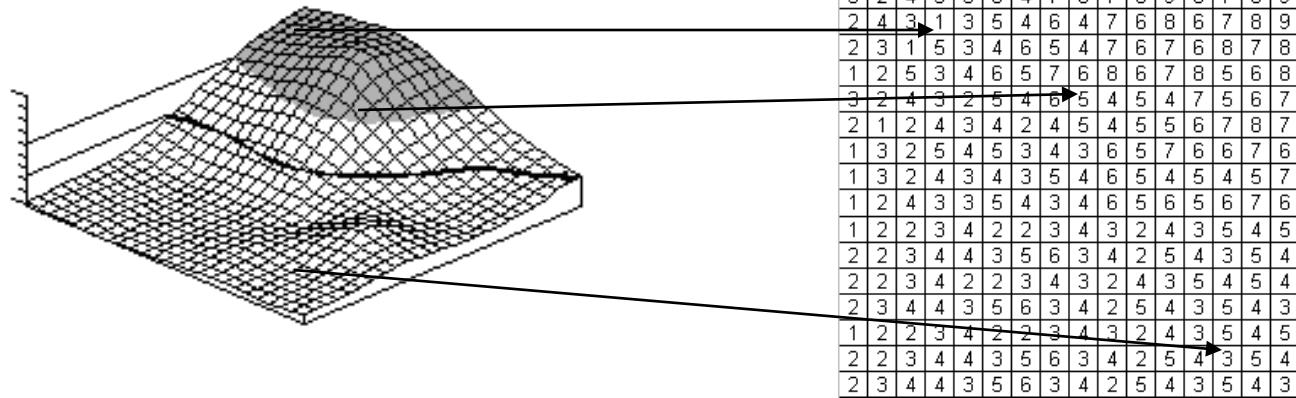


Vemos lo que esperamos ver



Comparando Mapas

Los mapas se transforman en una grilla de números

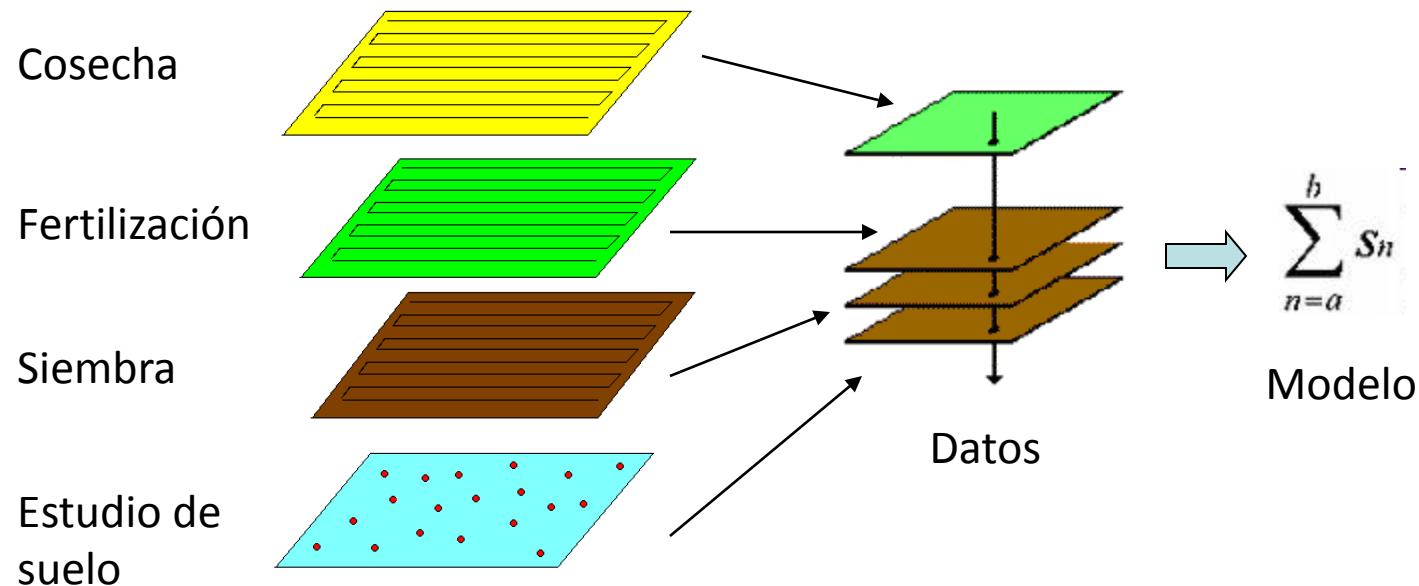


Ahora es posible usar análisis estadísticos para saber qué tan similares son dos mapas



Modelos de Predicción

Los modelos representan las relaciones entre las variables



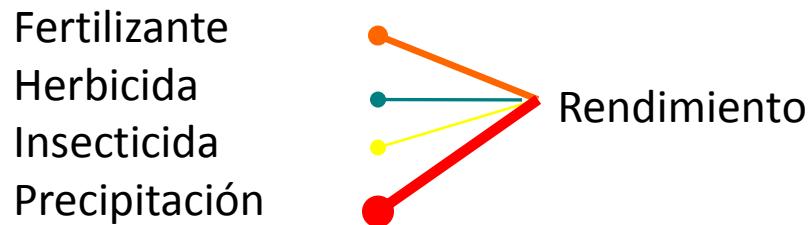


Modelos de Predicción

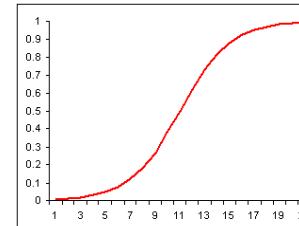
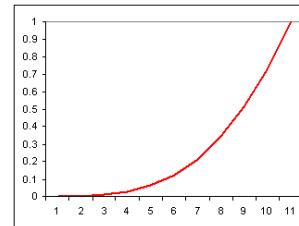
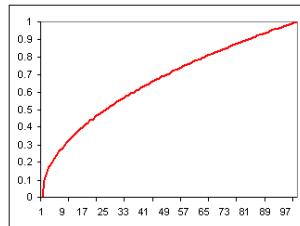
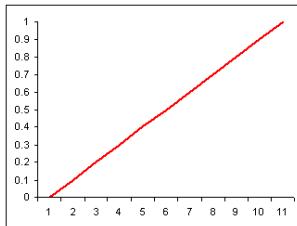
Los modelos representan relaciones entre las variables independientes y la variable dependiente

Responden a preguntas tales como:

¿Qué variable es más importante?



¿Qué forma tiene la relación?





Modelos para Controlar

Una vez que se entienden las relaciones entre las distintas variables es posible usar los modelos para controlar un resultado mediante la generación de prescripciones



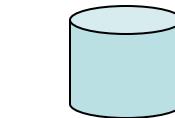
Proceso de Mejor Continua



GPS



Trabajo con
Tecnología GPS



Captura de
datos



Análisis de datos para
extraer información



Toma de decisiones



Etapas

- ✓ Caracterizar lotes en base al rendimiento
- ✓ Descubrir relaciones entre rendimiento y variables disponibles
- ✓ Diseñar ensayos para confirmar las relaciones encontradas y descubrir otras



Intercambio

¿Preguntas?