



Triunfa con BIGDATA



Frank Mauricio Rigal
Ingeniero Senior Líder en Proyectos
de Network Analytics en Telefónica

Un poco Sobre la Persona que quiere Compartir lo Aprendido con Ustedes:

1. Profesor Instructor de Pregrado en materias de Arquitectura del Computador, Programación de MicroControladores, Algoritmo de Máquinas de Estado, Electrónica Digital, matemáticas y geometría analítica desde 2012 al 2016.
2. Desarrollador Web de Portales de Seguridad Laboral, Programas de Lealtad, Ecommerce, Tipificaciones de CallCenter y Plataformas de Monitoreo IOT en empresas internacionales y StarTups Desde el 2011 al 2014 y 2016.
3. Coordinador Técnico de Analítica de Canales en Telefónica Venezuela, Datawarehouse de Onnicanalidad, Migración de Modelo de Datos del CallCenter, Generación del Modelo de Llamadas Salientes de CallCenter desde el 2014 al 2016.
4. Ingeniero en Proyectos de Network Analytics en Telefónica desde el 2016 a la actualidad, Modelo centralizado de Red BAF, Detección de Fallas Masivas FTTX.
5. Maestría en Gerencia Empresarial cursada desde el 2014 al 2016.
6. Diplomado en Bigdata con foco en arquitectura titulado en el 2016 en UTCV.
7. Diplomado en Bigdata con foco en Ciencia de Datos titulado en el 2017 en la UC.
8. Maestría Telefónica en Bigdata y Business Analytics titulado en 2018 por UEMC.
9. Scrum Master Certified desde el 2016.
10. Actualmente cursando MBA con core en Negocios e Innovación en la UDD.

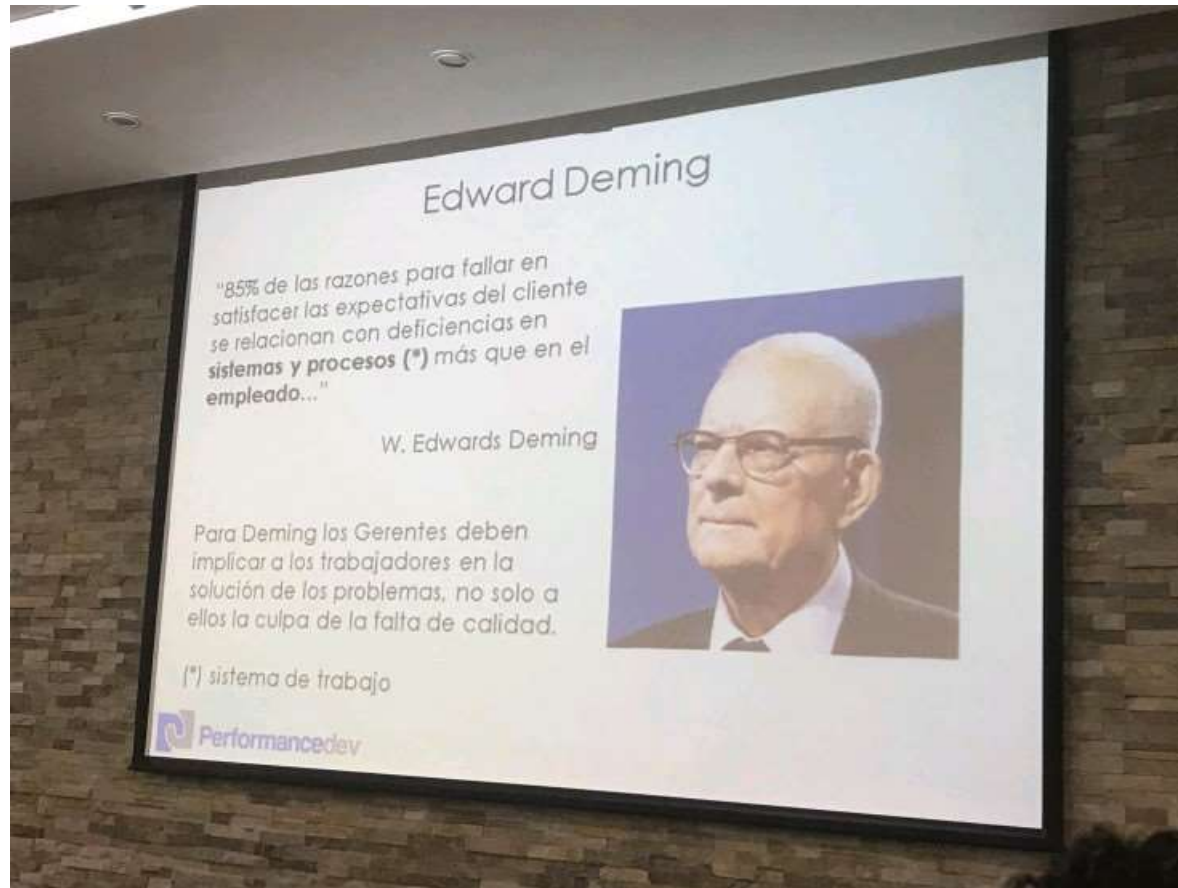
Productivizar el Bigdata

Las organizaciones necesitan un proceso bien definido para llegar a integrar descubrimientos analíticos en los sistemas operacionales de las mismas.



BIG DATA WEEK
A GLOBAL FESTIVAL OF DATA

Productizar el Bigdata



Que nuestros Proyectos de Analítica de Datos tengan un Propósito



Que nuestros Proyectos de Analítica de Datos tengan un Propósito

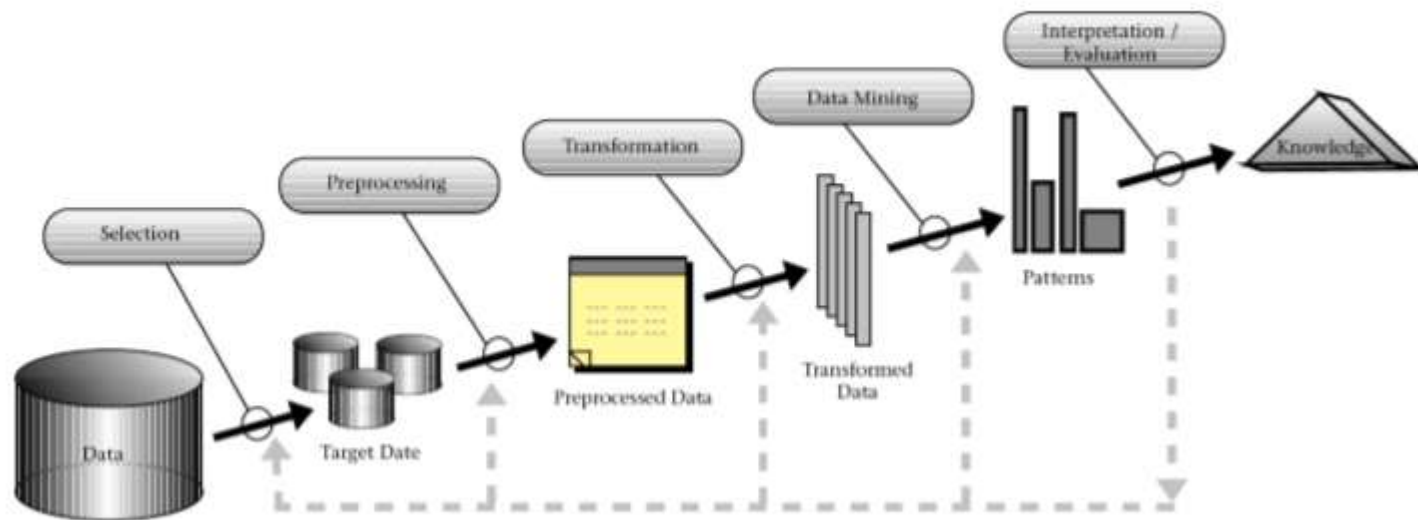


Mi equipo Ganador

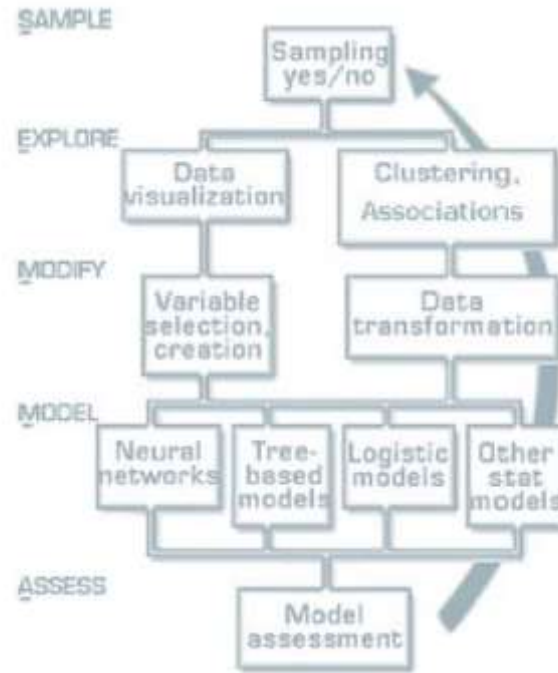


BIG DATA WEEK
A GLOBAL FESTIVAL OF DATA

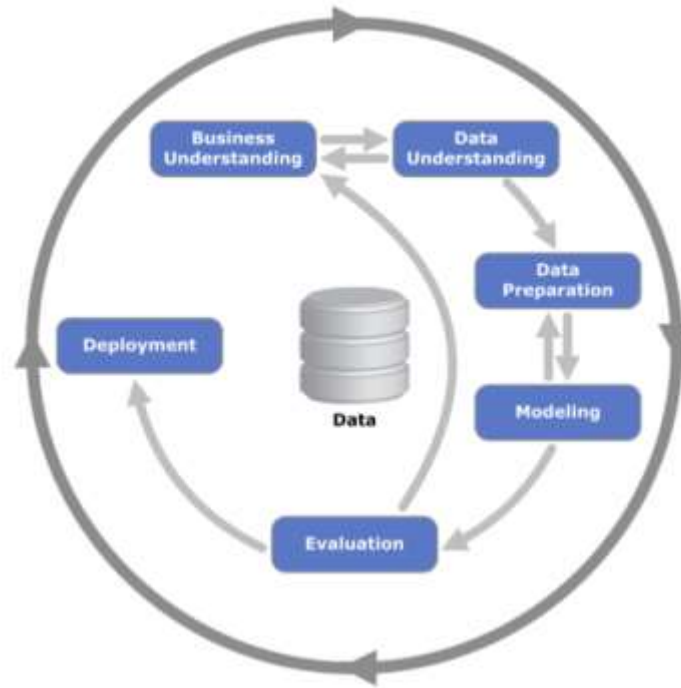
Metodologías de PDA KDD, SEMMA y CRISP-DM



**KDD: Proceso centrado en la extracción de conocimiento de un repositorio.
(Knowledge Discover Database)**



SEMMA: Proceso centrado en la minería de datos. (Sample, Explore, Modify, Model and Assess)



CRISP-DM: Proceso iterativo y centrado en el negocio (Cross Industry Standard Process for Data Mining)

Comparativa de las 3 metodologías

KDD	SEMMA	CRISP-DM
▪ Pre KDD	▪ -----	▪ Business understanding
▪ Selection	▪ Sample	▪ Data understanding
▪ Pre processing	▪ Explore	▪ Data preparation
▪ Transformation	▪ Modify	▪ Modelling
▪ Data mining	▪ Model	▪ Evaluation
▪ Interpretation/Evaluation	▪ Assesment	▪ Deployment
▪ Post KDD	▪ -----	

KDD	SEMMA	CRISP-DM
▪ Pre KDD	▪ -----	▪ Business understanding
▪ Selection	▪ Sample	▪ Data understanding
▪ Pre processing	▪ Explore	▪ Data preparation
▪ Transformation	▪ Modify	▪ Modelling
▪ Data mining	▪ Model	▪ Evaluation
▪ Interpretation/Evaluation	▪ Assesment	▪ Deployment
▪ Post KDD	▪ -----	

En esta sección detallamos cada etapa del proceso de Análisis de Datos, siguiendo fundamentalmente CRIPS-DM, pero deteniéndonos algo más en las fases previas de pre-procesado de los datos.

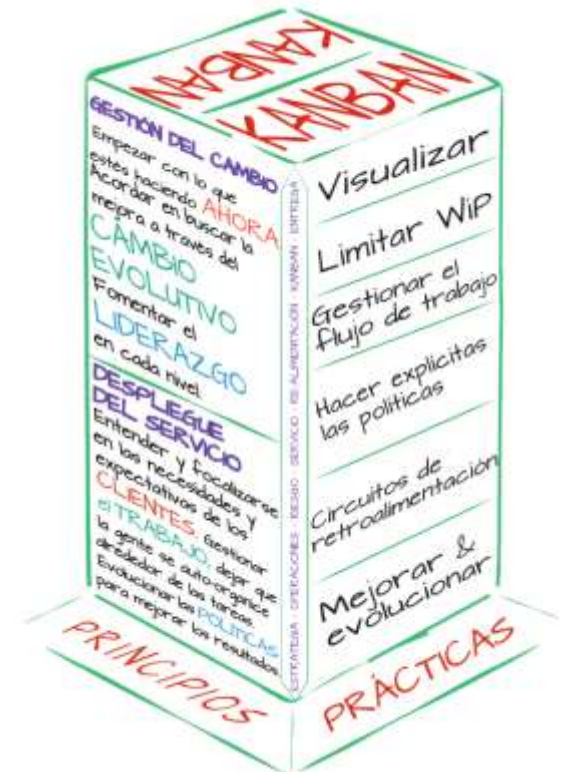
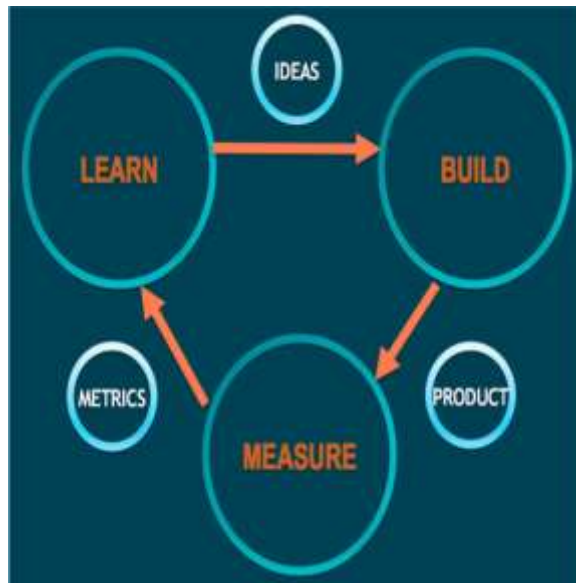
- Entendimiento del Negocio
- Definición del problema
- Definición de los datos
- Entendimiento de los datos
- Recolección de los datos
- Calidad de los datos
- Estadística descriptiva
- Preparación de los datos
- Muestreo
- Selección y Transformación de variables
- Adaptación de variables
- Modelado
- Evaluación
- Interpretación de resultados
- Síntesis de resultados
- Despliegue



BIG DATA WEEK
A GLOBAL FESTIVAL OF DATA

Fuente: <http://recipp.ipp.pt/bitstream/10400.22/136/1/KDD-CRISP-SEMMA.pdf>

Entendimiento del Negocio al Modelado Con Design Thinking



BIG DATA WEEK
A GLOBAL FESTIVAL OF DATA

Data Journalism Handbook

¿Por qué fracasan los Proyectos de Bigdata?

Causas de fracaso de proyectos Big Data



Tamaño organización

- <1000 empleados
- 1000+ empleados



BIG DATA WEEK
A GLOBAL FESTIVAL OF DATA

Metodología Agile ScrumBan



BIG DATA WEEK
A GLOBAL FESTIVAL OF DATA

Metodología Agile ScrumBan

¿Qué causa que los proyectos ágiles fallen?



Source - VersionOne

Despliegue



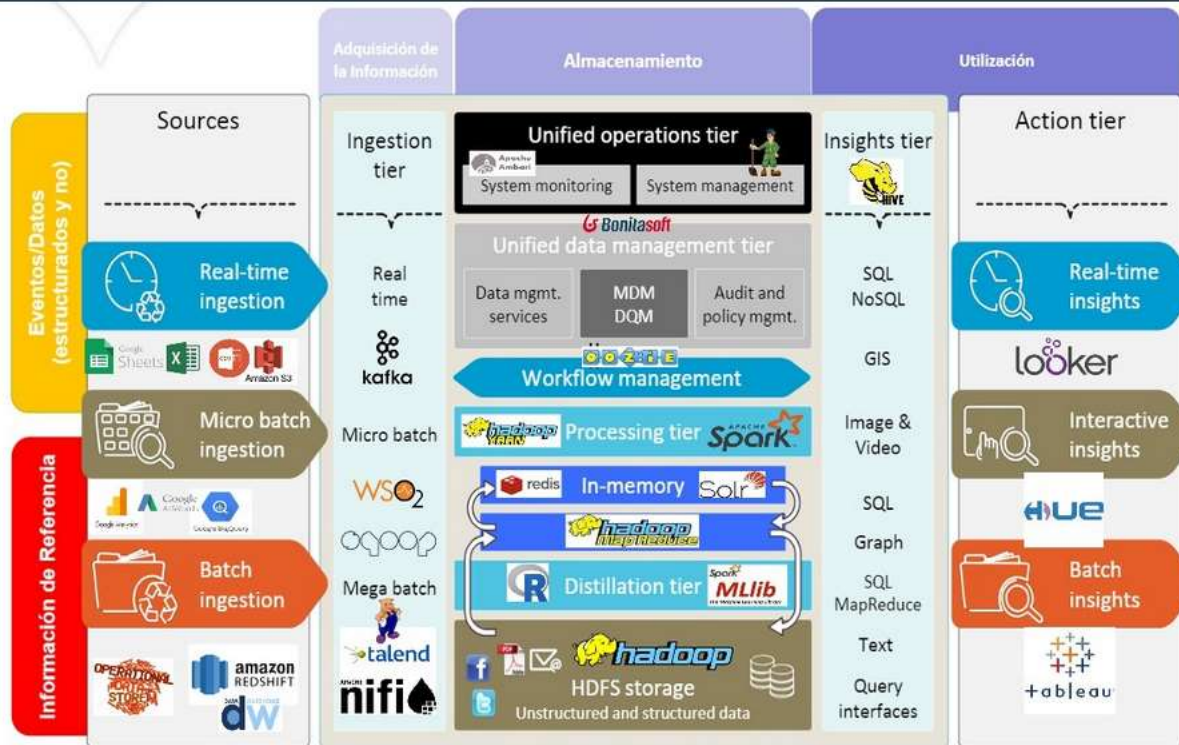
Amazon
Web Services



Microsoft
Azure

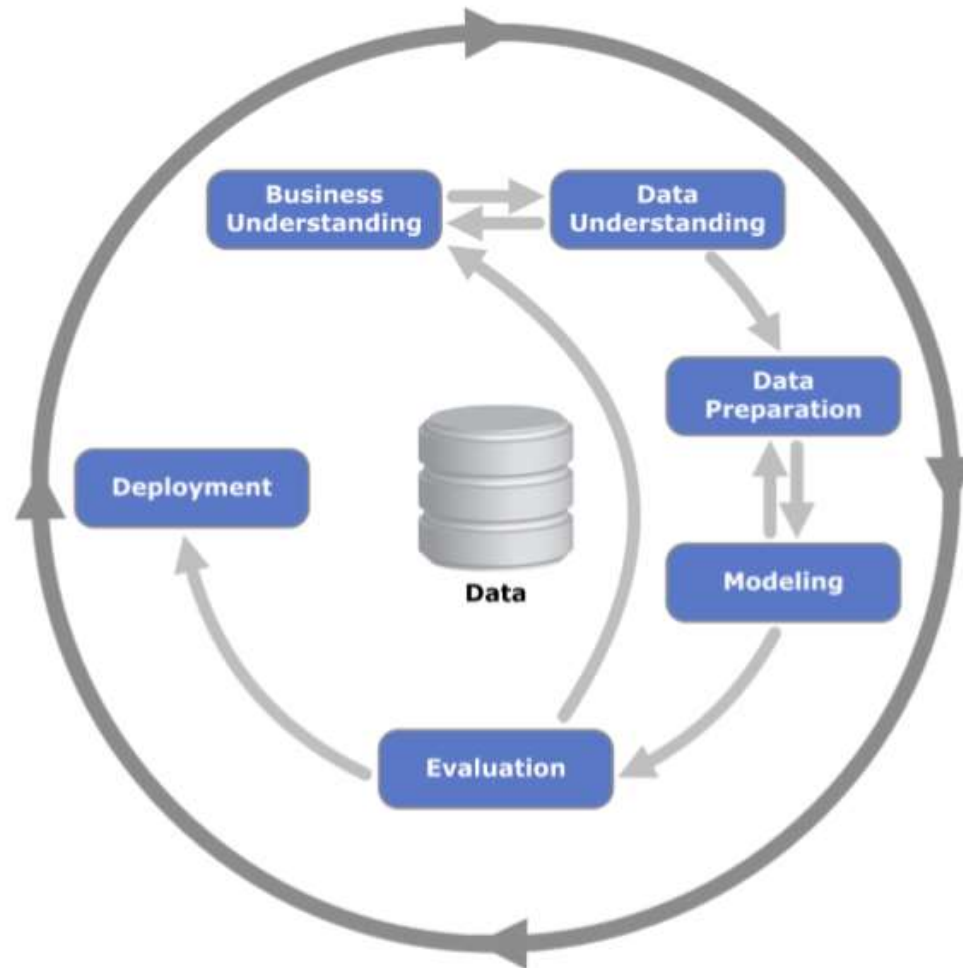


Google
Cloud

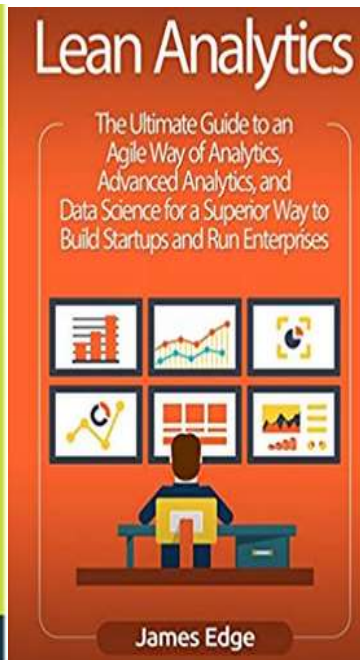
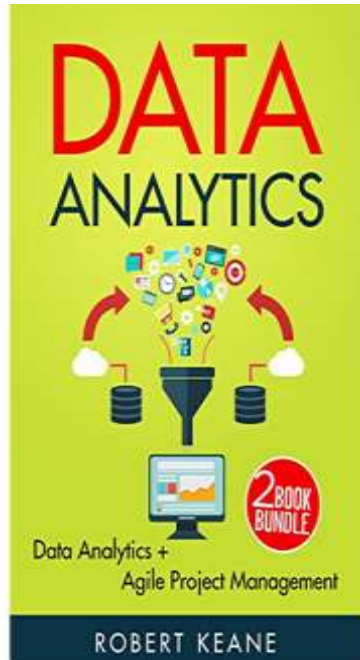
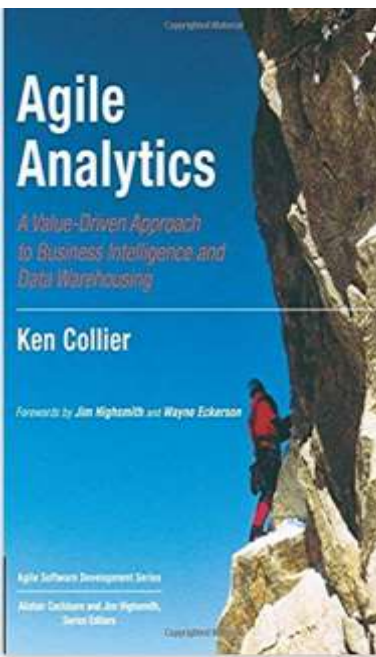


Flujo Continuo de Disminución de Incertidumbre

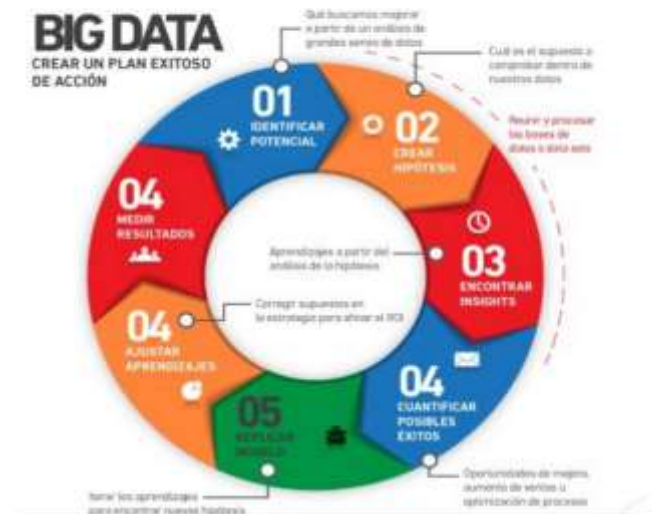




Tendencias



De esta forma, la metodología propuesta para llevar a cabo los proyectos de Big Data es **Scrum** (metodología ágil y flexible para gestionar el desarrollo de software), con 4 **Sprints**; los dos primeros de 2 semanas; el tercero de 4 semanas; y el último de 6 semanas; para un total de 12 semanas, con revisiones diarias (**daily**). (ver, Figura 1).



Ciencia de Datos ágil

1 h 59 m • Empieza ahora



<https://www.linkedin.com/pulse/metodolog%C3%ADa-de-implementaci%C3%B3n-la-arquitectura-pedro-bonillo-ramos/>

Preguntas



BIG DATA WEEK
A GLOBAL FESTIVAL OF DATA