

Fundamentos de Ciencias de Datos

Plan del Curso y Gestión de Datos

Sesión 1

Agenda

- Conociéndonos
- Presentación del curso
 - Descripción
 - Resultados de aprendizaje
 - Estructura y contenido didáctico
 - Ponderación
 - Descripción de actividades del curso
- M1: Introducción a Ciencia de Datos

Conociéndonos



Fundamentos de Ciencia de Datos

Descripción

- La información digital ha redefinido la forma en la que las organizaciones se enfrentan a la explotación de los datos para mejorar su posición competitiva y tomar decisiones.
- Con un adecuado análisis de Ciencia de Datos se pueden:
 - detectar tendencias
 - realizar predicción de sucesos futuros
 - extraer patrones del comportamiento de los usuarios adaptando mejor los servicios a sus necesidades

Fundamentos de Ciencia de Datos

Resultados del aprendizaje

Los estudiantes serán capaces de:

- Saber qué es Ciencia de Datos, cuáles son sus orígenes, sus principales fuentes, los retos a los que se enfrenta y los principales tipos de aplicaciones que se pueden llevar a cabo.
- Determinar cómo y cuándo surgen éstos retos en diferentes campos de actividad.
- Utilizar herramientas de Ciencia de Datos.
- Identificar las consecuencias que la generalización del uso de Ciencia de Datos está teniendo en las organizaciones.

Fundamentos de Ciencia de Datos

Estructura y contenidos didácticos

El curso está estructurado en cuatro módulos didácticos, el seguimiento de un caso práctico de estudio, un proyecto final integrador, actividades de reforzamiento de los contenidos y un examen final.

Los módulos que integran el curso son los siguientes:

M1 –Introducción a Ciencia de Datos

M2 –Trabajo con datos

M3 –Gestión de los datos

M4 –Visualización y consumo de los datos

Fundamentos de Ciencia de Datos

Ponderación:

Reportes de aprendizaje (ClickStream)	20%
Proyecto final	40%
Exámenes rápidos, control de lecturas, actividades	10%
Examen final	20%
Participación en clase	10%
Calificación Final	100

Fundamentos de Ciencia de Datos

Características de las reportes (ClickStream)

Durante la elaboración de los reportes, ten en consideración los siguientes aspectos:

- Elabora una plantilla propia para la entrega de todos los reportes.
- Se desarrollan de forma individual.
- Se deben cubrir las especificaciones de cada entregable.
- Calidad en la redacción y estructura del reporte.
- Subirlos a BB en tiempo y forma. No se aceptan reportes por correo electrónico, ni fuera de tiempo.
- Aportación obtenida del reporte.

Fundamentos de Ciencia de Datos

Características del Proyecto Final

- El proyecto final consistirá en la réplica de los reportes (ClickStream) en un área de aplicación específica seleccionada por los estudiantes (salud, gobierno, mercadotecnia, finanzas, educación, etc.)
- Se desarrollará en equipos de tres personas.
- Lo ideal es contar con una empresa, o seleccionar un giro de organización que sea posible importar o bajar los archivos con datos de Internet.
- Debe incluir todas las actividades elaboradas en los reportes.
- Asegurar la calidad en la redacción y estructura de los entregables.

Fundamentos de Ciencia de Datos

Exámenes rápidos , control de lecturas y actividades

- Se deben cubrir las especificaciones de cada entregable.
- Las actividades deben ser entregadas en tiempo y forma en BB.
- Cualquier trabajo de investigación o lectura podrá ser motivo de examen rápido.
- Calidad en la redacción y estructura de las actividades.
- Las investigaciones deben incluir mínimo seis referencias bibliográficas.

Fundamentos de Ciencia de Datos

Examen final

- Examen escrito.
- De manera presencial y síncrona.
- Incluye todo el material visto durante el curso.
- Se realiza en la fecha establecida en el plan del curso:
 - 27 de abril de 2021.

Fundamentos de Ciencia de Datos

Participación en clase

- Se llevará un registro de las participaciones en clase a través de diversas herramientas.

Participación 01: Individualmente comentar **una** noticia del área de Ciencia de Datos. La pueden exponer durante todo el cuatrimestre, excepto en las sesiones de presentaciones de avances y examen final, la duración de cada intervención no debe exceder los cinco minutos. Se mostrarán máximo dos intervenciones por sesión.

Fundamentos de Ciencia de Datos

Consideración:

- Es importante tener en cuenta que por la naturaleza del curso y los temas a cubrir, será necesario utilizar algunas herramientas en línea, y otras de instalación en su equipo de trabajo.



Fundamentos de Ciencias de Datos

M1 - Introducción a Ciencia de Datos

Sesión 1

M1 – Introducción a Ciencia de Datos

Introducción

El módulo 1 se centra en el análisis de los tres puntos siguientes:

- La gestión de los datos.
- El concepto de Ciencia de Datos
- El rol que la Ciencia de Datos desempeña en las organizaciones.

M1 – Introducción a Ciencia de Datos

Objetivo y resultados de aprendizaje

Objetivo

- Presentar qué se entiende por gestión de datos.

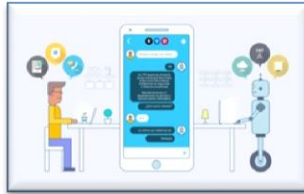
Resultados del aprendizaje

- Identificar la evolución de los sistemas de gestión y análisis de datos.
- Conocer qué son los sistemas operacionales y los sistemas de apoyo a la toma de decisiones.
- Reconocer los fundamentos del trabajo de datos.
- Saber cuál es el ciclo típico de vida de los datos.

M1 – La Gestión de los Datos

Evolución en los Sistemas de Gestión y Análisis de Datos

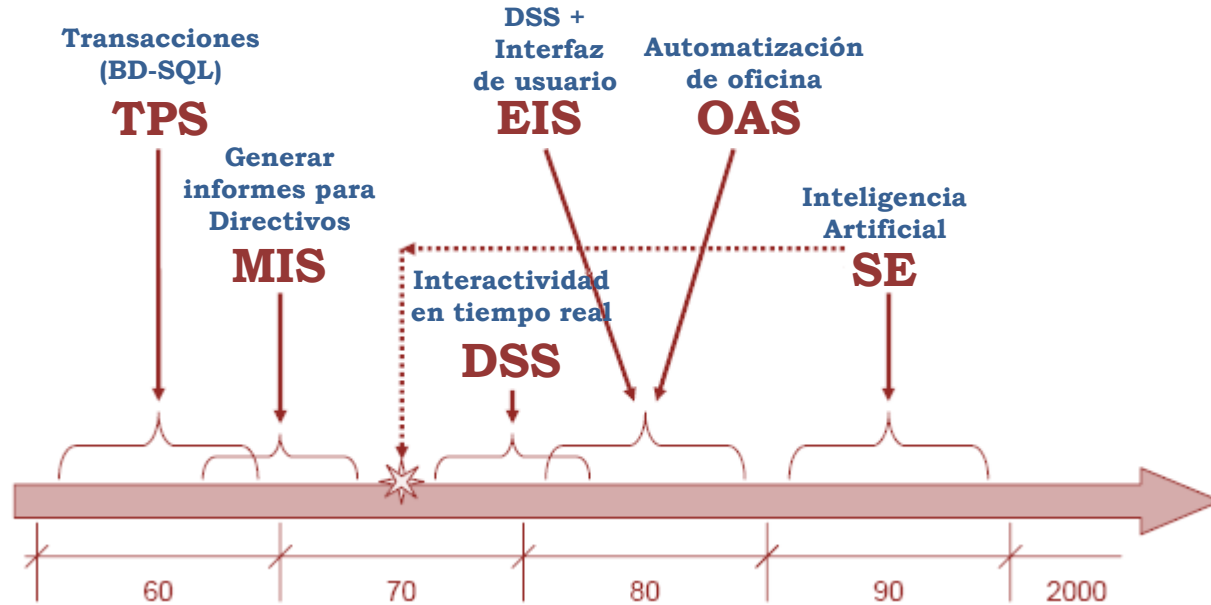
- Factores que influyen:
 - Aparición y consolidación de nuevos avances tecnológicos.



- Continuo abaratamiento de los costes asociados a éstas tecnologías.

M1 – La Gestión de los Datos

Evolución en los Sistemas de Gestión y Análisis de Datos



M1 – La Gestión de los Datos

Evolución en los Sistemas de Gestión y Análisis de Datos

Procesamiento de transacciones en línea (OLTP)	Procesamiento analítico en línea (OLAP)
Almacén de datos actuales	Almacén de datos históricos
Almacena datos al detalle	Almacena datos al detalle y datos agregados a distintos niveles
Bases de datos medianas	Bases de datos grandes
Los datos son actuales	Los datos son estáticos
Estructura relacional	Estructura multidimensional
Los procesos son repetitivos	Los procesos no son previsibles
Tiempo de respuesta pequeño	Tiempo de respuesta variable
Soporta decisiones diarias	Soporta decisiones estratégicas

M1 – La Gestión de los Datos

Cada minuto en Internet suceden:



4,1 millones de
búsquedas en Google



3,3 millones de
actualizaciones en
Facebook



347,000
tweets



347.222 *stories*
subidas a
Instagram



10 millones de
anuncios
visualizados



+500 horas de
vídeo subidas a
YouTube

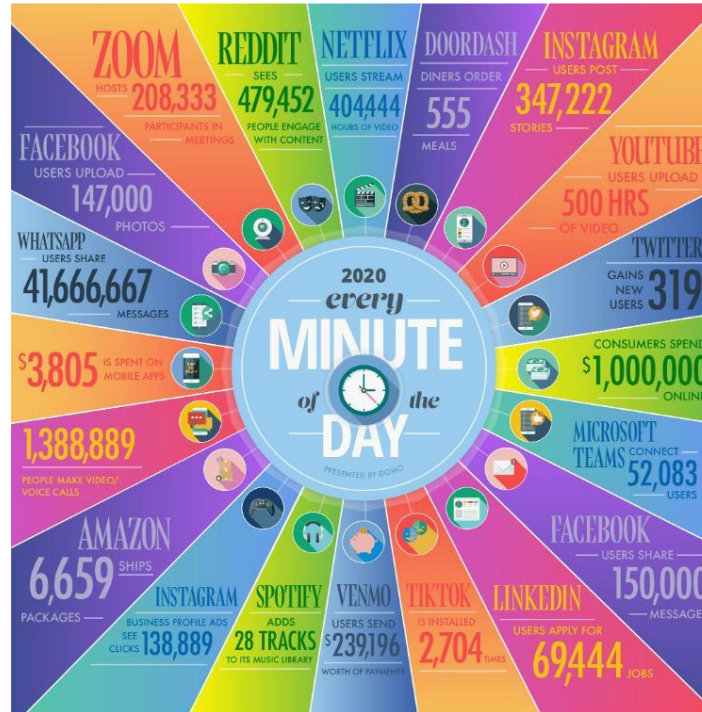


32,000 horas
de música en
streaming



194,000
aplicaciones

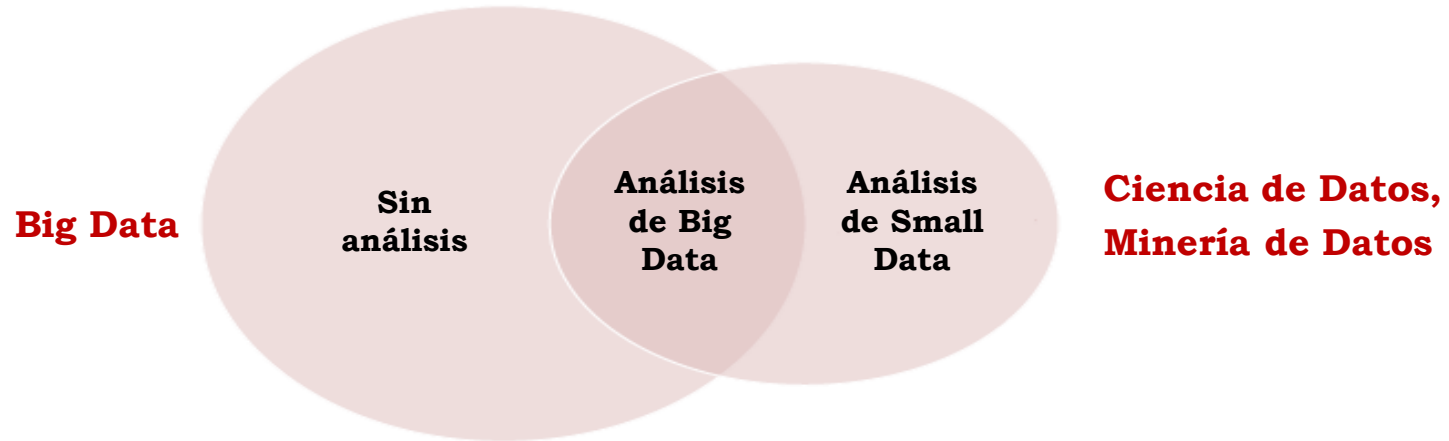
M1 – La Gestión de los Datos



Fuente: <https://www.domo.com/learn/data-never-sleeps-8>

M1 – La Gestión de los Datos

Evolución en los Sistemas de Gestión y Análisis de Datos



No toda la **Ciencia de Datos** es **Big Data**.
No todo el **Big Data** es **Ciencia de Datos**.

M1 – La Gestión de los Datos

Acercándonos a la realidad



John Cohn: "en los próximos años los objetos serán inteligentes por estar conectados".

Participación 02 FORO

1. ¿Creen justificada su afirmación?
2. ¿En qué sentido Internet de las cosas puede contribuir a mejorar temas como la salud o el tráfico vehicular?

M1 – La Gestión de los Datos

Actividad 1

Nombre: Actividad 1: Del bit... al Big Data

Tipo: En equipos del proyecto final.

Descripción: Elabora un **mapa conceptual** del e-book de la empresa PowerData, Del bit... al Big Data. Documento anexo en BB.

Condiciones de satisfacción: El mapa conceptual debe mencionar todos los puntos tratados en el e-book.

Fecha de entrega: El día martes 26 de enero de 2021 en la hora de clase.

Forma de entrega: Archivo anexo en BB, en esta ocasión, aunque sea una actividad por equipo, todos suban la actividad aunque esté replicada.

Cómo se evaluará: Es parte de la ponderación de las actividades de aprendizaje del curso.

M1 – Bibliografía

- ✓ Rowley, Jennifer (2007). «The wisdom hierarchy: representations of the DIKW hierarchy». Journal of Information Science 33 (2): 163-180. doi:10.1177/0165551506070706.
- ✓ McGahan, A., Unlocking the Big Promise of Big Data, En: Rotman Magazine, 2013.
- ✓ Sistema de Soporte a las Decisiones, https://www.ecured.cu/Sistema_de_Soporte_a_las_Decisiones, consultado en línea el 13 de enero de 2021.
- ✓ Dasu T., Johnson T., Exploratory Data Mining and Data Cleaning, Wiley, 2003.
- ✓ Fayyad, U., Shapiro, G., From Data Mining to Knowledge Discovery in Databases, En: AI Magazine 17.3 (1996), Págs. 37-54.
- ✓ <https://www.corporatecomplianceinsights.com/mayer-browns-tech-talks-episode-3-the-big-data-paradox/>, consultado en línea el 13 de enero de 2021.
- ✓ <https://www.akamai.com/es/es/products/network-operator/dnsi-big-data-connector.jsp>, consultado en línea el 13 de enero de 2021.