

# Clase 3: Organizaciones & Data Driven.

Diplomado de Ciencia de Datos

Notas por José Daniel Conejeros - jdconejeros@uc.cl

Junio del 2020

**Profesor:** Francisco Kuncar - [fkuncar@gmail.com](mailto:fkuncar@gmail.com)

## I. Clientes, datos e información

En general las organizaciones se manejan (o se deberían manejar) a partir de los datos y la información que pueden obtener de sus clientes. En consecuencia, es necesario explorar, analizar y comprender a los clientes pues son **el principal activo de una empresa**. Pero, ¿quiénes son los clientes?

Hasta el año 2013 los clientes eran percibidos de forma muy homogénea, sin embargo, los procesos de cambio social han dado paso a un nuevo tipo de cliente exclusivamente egocéntrico. En ese sentido, la relación entre cliente y empresa se observa un poco más balanceada respecto al pasado y ahora los primeros son directos al exigir sus derechos hacia los segundos.

### a. ¿Qué nos interesa saber de los clientes?

Lo que debiera ser preguntado son los atributos que causan dolor en los clientes, los temas estructurales que afectan directamente a los clientes. En ese sentido se debe tener en cuenta:

- Hay una brecha entre lo pensado por la empresa y lo que el cliente ve. Una distancia entre lo planificado por la empresa y lo observado empíricamente por el cliente.
- Se deben manejar **las expectativas** pues inciden en lo que cada uno quiere. Si no tenemos la capacidad de responder a las expectativas, no podremos mantener a nuestros clientes.

### b. ¿De qué nos sirven los datos?

Los datos e indicadores nos ayudan a generar una estrategia de expectativas sólidas. En la medida de que se cuente con **atributos medibles** de los clientes, se podrá trazar una estrategia de desarrollo capaz de captar los problemas y resolverlos. Se debe ser proactiva para anticiparse a distintos escenarios a través de un sistema de predicción y alerta.

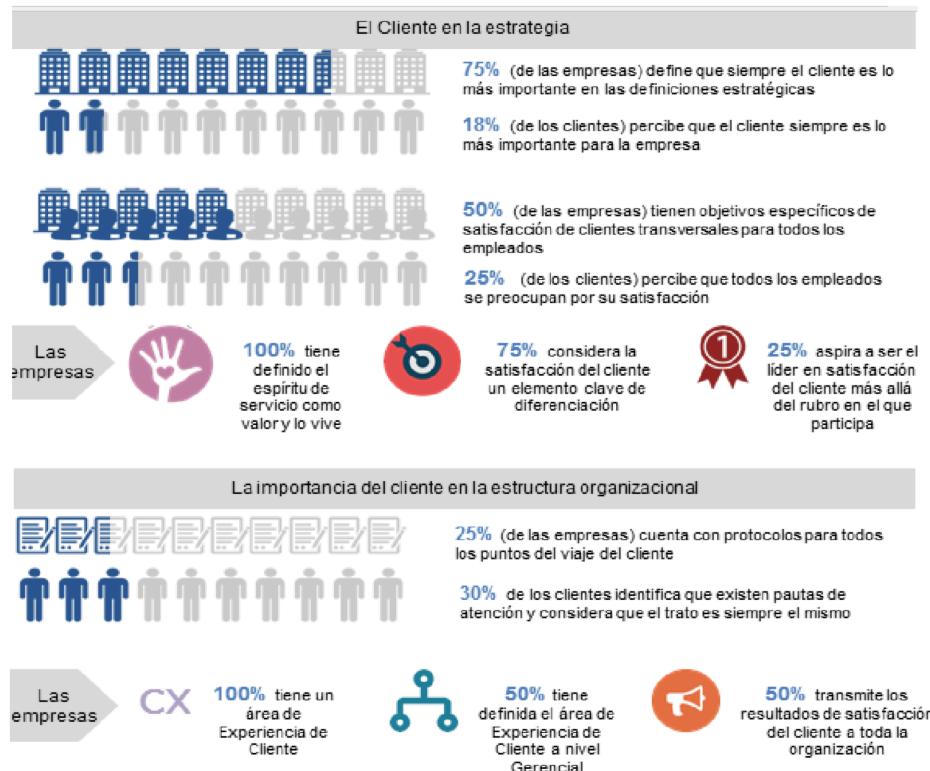


Figura 1: Data sobre clientes

Con los datos a disposición podemos diagnosticar y generar acciones buenas y concretas para responder a los problemas y mientras mejor sea el enganche con el cliente, mayor propensión tendremos a que generen insumos informativos.

## II. La digitalización

El cambio en los problemas y las prácticas cotidianas de los clientes condicionan los servicios de las empresas. Bajo ese concepto emergen organizaciones que responden a necesidades o problemáticas de la vida cotidiana de manera innovadora y flexible respecto a la oferta existente. Estas nuevas organizaciones (uber, netflix, spotify, etc.) responden a este desafío y fuerzan a generar más competencia aumentando la calidad de sus servicios. La clave es dejar la tecnología a la disposición de los clientes:

- Nuevas formas de entregar bienes y servicios
- Nuevas formas de distribución
- Nuevas formas y dinámicas de trabajo
- Nuevos desafíos de mercado

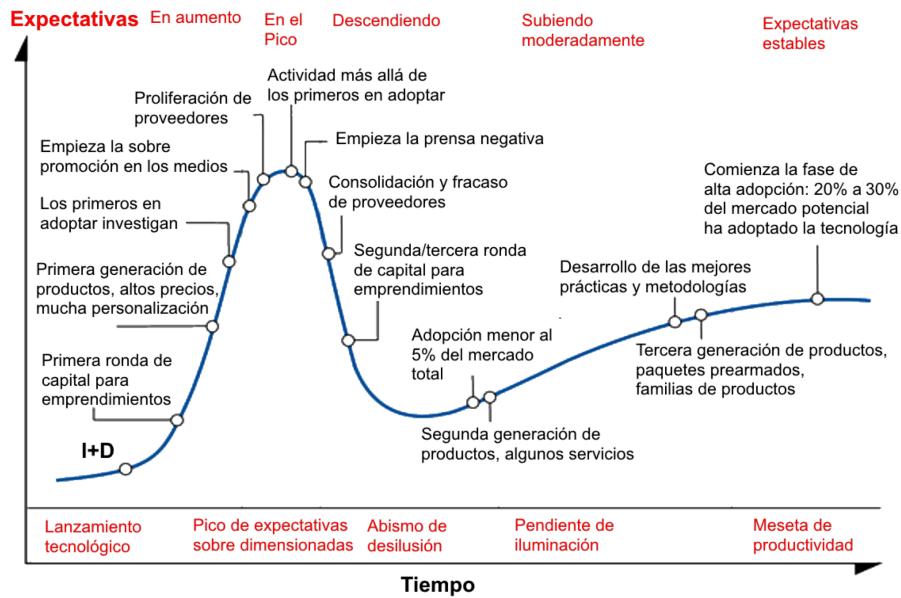


Figura 2: Productividad y expectativas

### a. El dato

A partir del 2005 las empresas empezaron a acumular datos, información, pero sin un desarrollo tecnológico capaz de canalizarla (solo con el fin de almacenarla). No obstante, con el tiempo se comenzó a considerar la riqueza propia de los datos y se necesito de analistas capaces de explotar esa información, de ver cosas que el resto no ve. En consecuencia, fue necesario un **mayor desarrollo de las empresas en tecnología** para explorar, analizar y comprender sus datos. En la actualidad los datos están en nuestra vida cotidiana al estar permanentemente conectados a través de aparatos móviles:

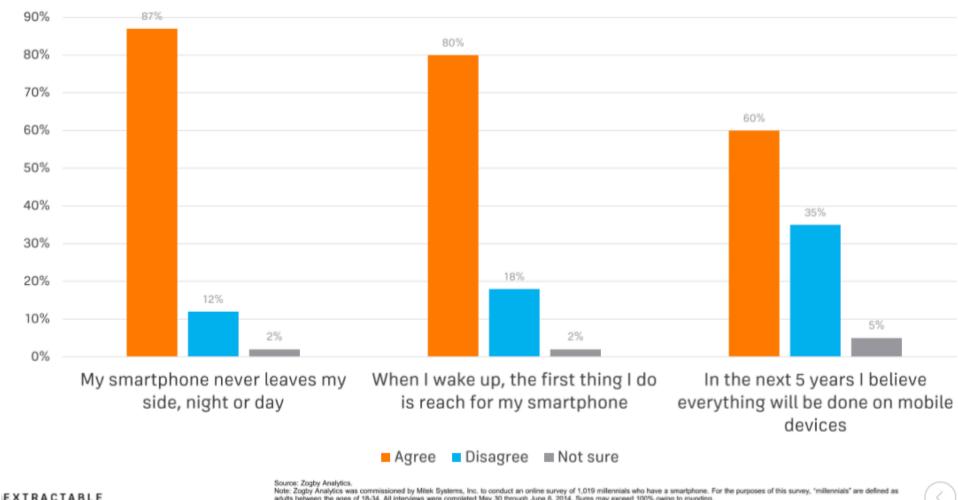


Figura 3: Uso de móviles

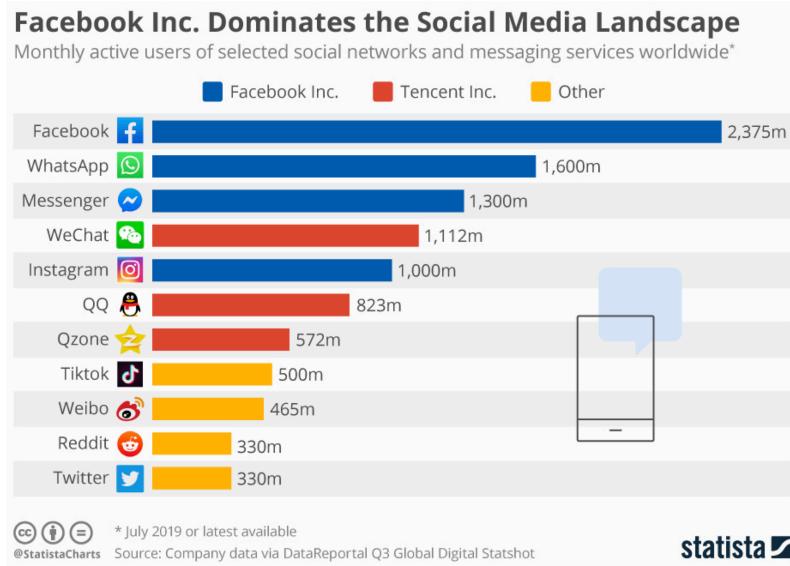


Figura 4: Usuarios activos en redes sociales

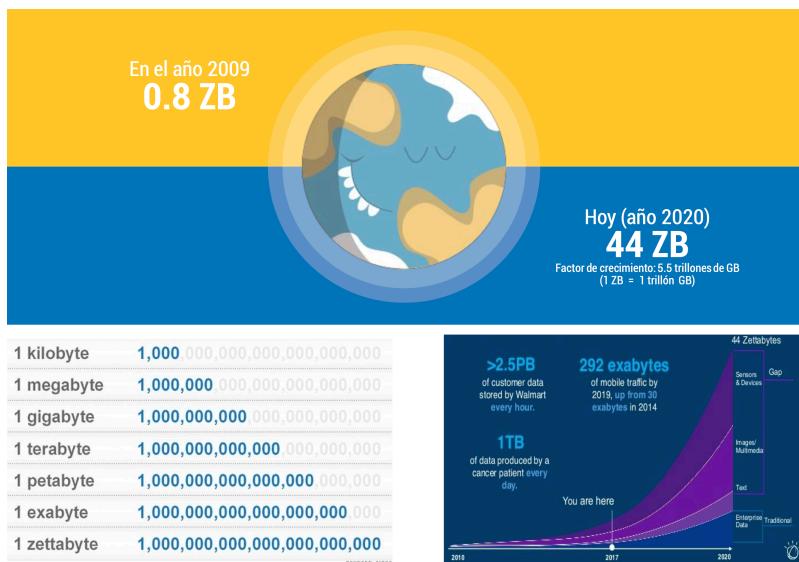


Figura 5: Información en el mundo

### b. Análisis para generar valor

El posicionamiento de una organización puede ser guiada por los datos. En la actualidad hay un alto volumen de información generada por los/las usuarios y a pesar de que tenemos muchos, seguimos ocupando muy poca información en un mundo que es dinámico, cambiante y exigente. Se requieren herramientas de análisis en tiempo real y con diversos tipos/fuentes de datos.

Se debe **transformar los datos en activos de valor** para que tomar decisiones correctas en base a hechos empíricos. En definitiva apoyar una toma de decisiones buscando valor.



Figura 6: Modelos de análisis

Para lograr esta transformación se requiere un cambio de paradigma respecto a la gestión de información y procesos de gestión. En términos sencillos es el paso de un proceso manual a procesos automatizados que generen valor:

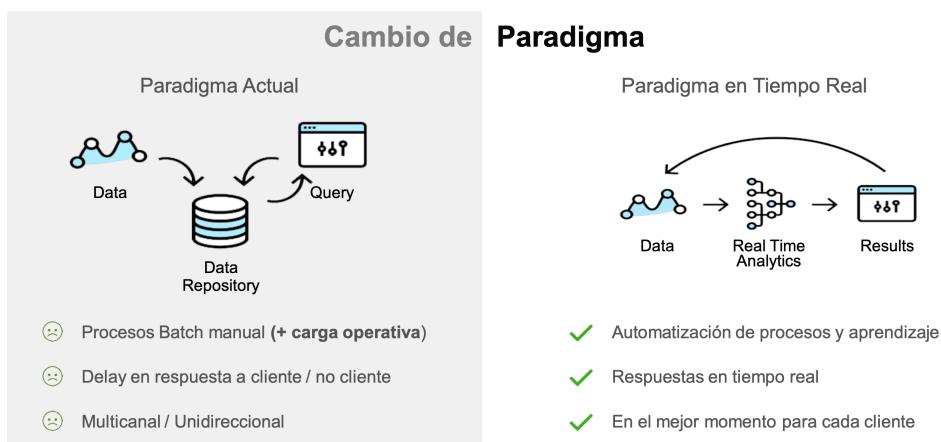


Figura 7: Paradigmas de gestión

## II. Proyectos de Big Data

Las claves de un proceso de Big Data se centran en la gestión de datos, generación de insights a partir de esa información y las acciones tomadas (en base al objetivo del negocio):

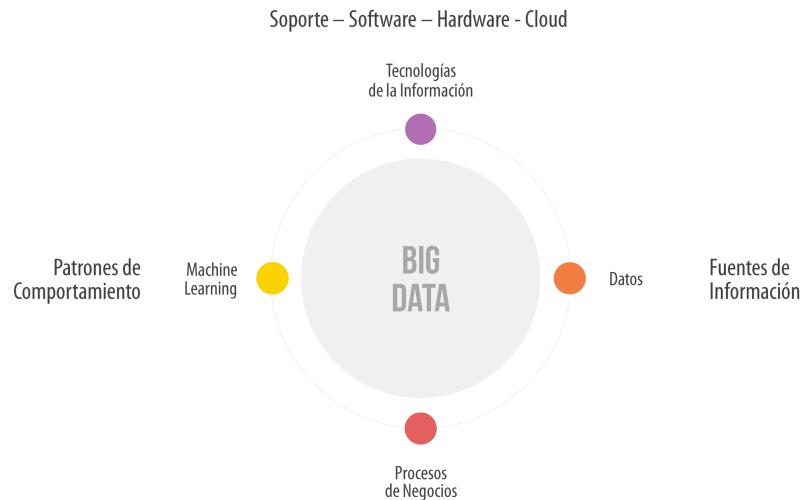


Figura 8: Proceso de Big Data

- Tecnología de información que tengo a disposición.
- Datos de distintas fuentes son transformados en valor.
- Machine Learning: Necesito algoritmos para entender los patrones del comportamiento.
- Integrar estos resultados al proceso y decisiones del negocio.

Mucho dato hace que uno se vuelva reactivo. Analytics de alto nivel busca adelantarse a los posibles escenarios y esto necesita ser valorizado, **monetizado**. Se debe considerar que el valor de un proceso de análisis tiene una relación positiva con la capacidad analítica del mismo:

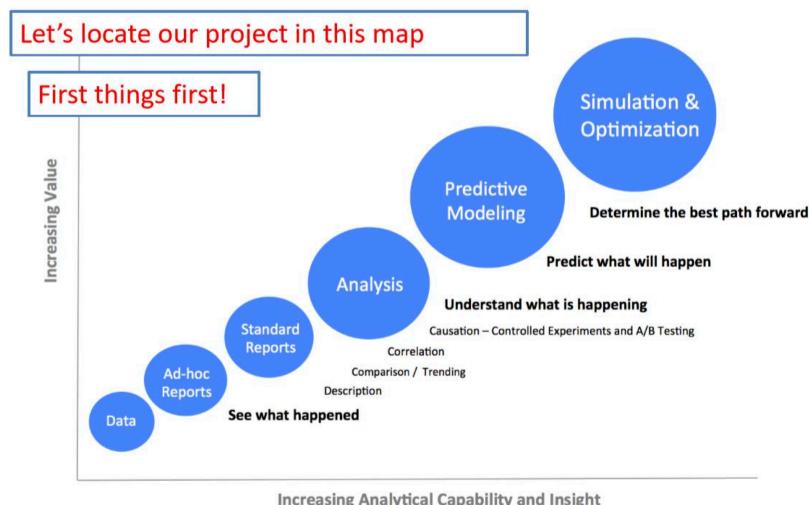


Figura 9: Valorización de análisis

Todo **buen proyecto de Big-Data** debería considerar la siguiente secuencia:

1. Definir objetivos
2. Definir datos
3. Saber que se puede obtener
4. Obtener los datos
5. Limpieza de los datos
6. Realizar análisis exploratorios
7. Modelamiento
8. Interpretación
9. Resultados
10. Reportes
11. Automatizar códigos
12. Distribuir y visualizar
13. Usar insights y acciones

En términos de **análisis** se consideran los siguientes:

- Análisis Descriptivos
- Análisis de Diagnóstico
- Análisis Predictivo
- Análisis Prescriptivo

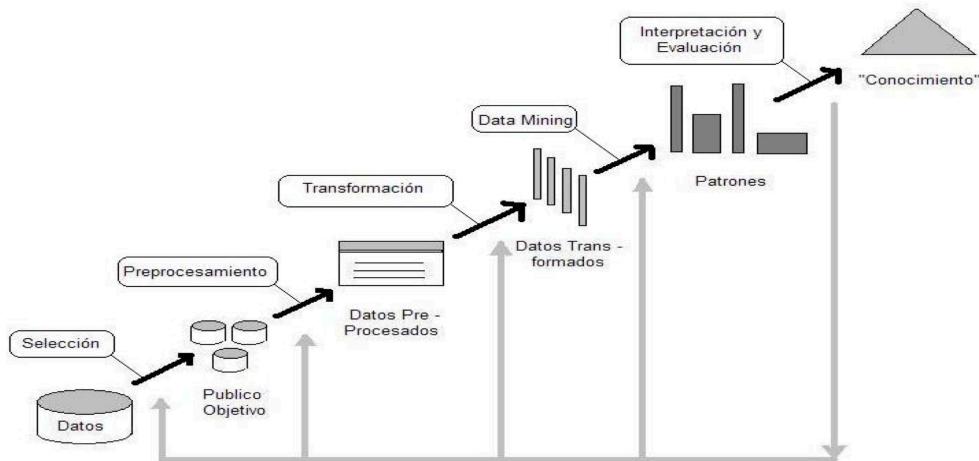


Figura 10: Aplicaciones de Machine Learning

A su vez, los modelos de **Machine Learning**:

- Descriptivos: se busca encontrar propiedades generales de una muestra y se descubren características más importantes.
- Predictivos: es fundamental el rendimiento del modelo cuando es aplicado a nuevos datos.
- Supervisados: requieren de conocimiento previo acerca de las clases a la que pertenecen los patrones. La mejor solución se encuentra en base a minimizar medidas de error entre el valor observado y el predicho por el algoritmo.
- No supervisados: la mejor solución se encuentra en base a la agrupación de los patrones de acuerdo a las medidas de similitud.

Y sus aplicaciones son muy variadas:

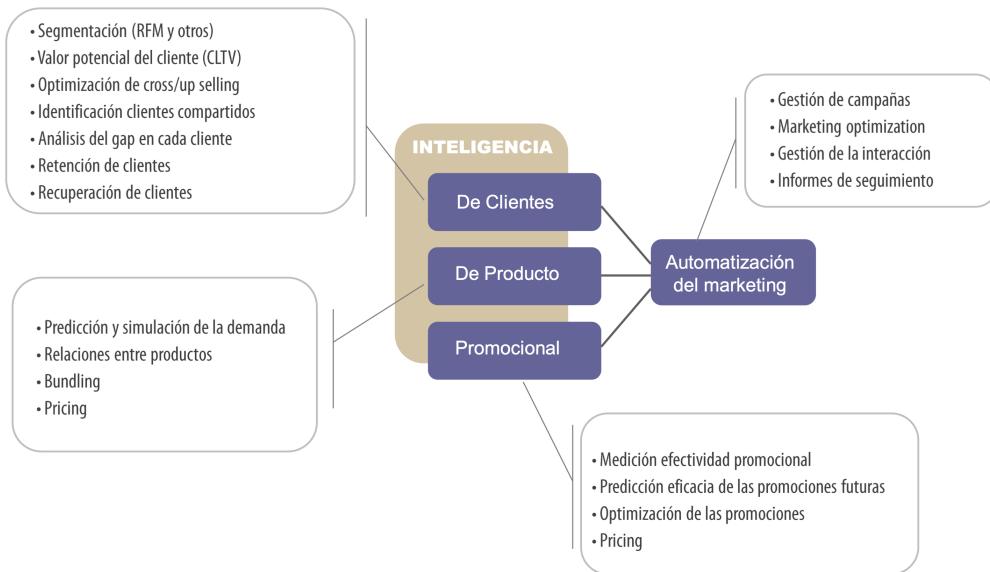


Figura 11: Aplicaciones de Machine Learning

Los algoritmos que existen para analizar conjuntos de datos:



Figura 12: Algoritmos de Machine Learning



Documento elaborado con las herramientas de [Rmarkdown](#)