

Transformación Digital

Ing. Luis Azevedo

Ing. Daniel Meerhoff

Ing. Quím. Ramiro Roselli

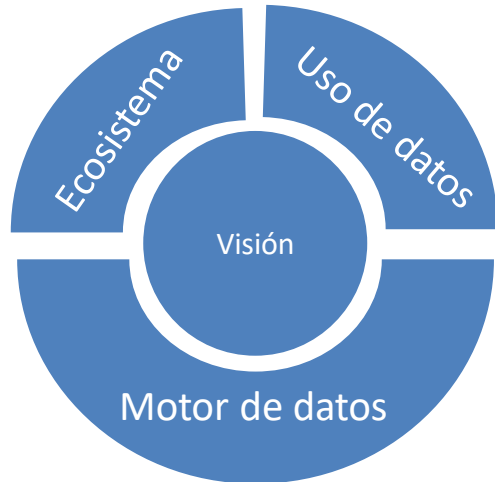


Contenidos

The background of the slide features a teal-to-white gradient. On the right side, there is a graphic with a silhouette of a person climbing a ladder towards a bright light. Surrounding this are various technical and business-related terms: 'INNOVATION' at the top, 'POWER' and 'ILLUMINATION' in the upper right, 'TEAM' on the left, and 'ENGINEERING' at the bottom. Binary code (0s and 1s) is scattered throughout the right side of the graphic.

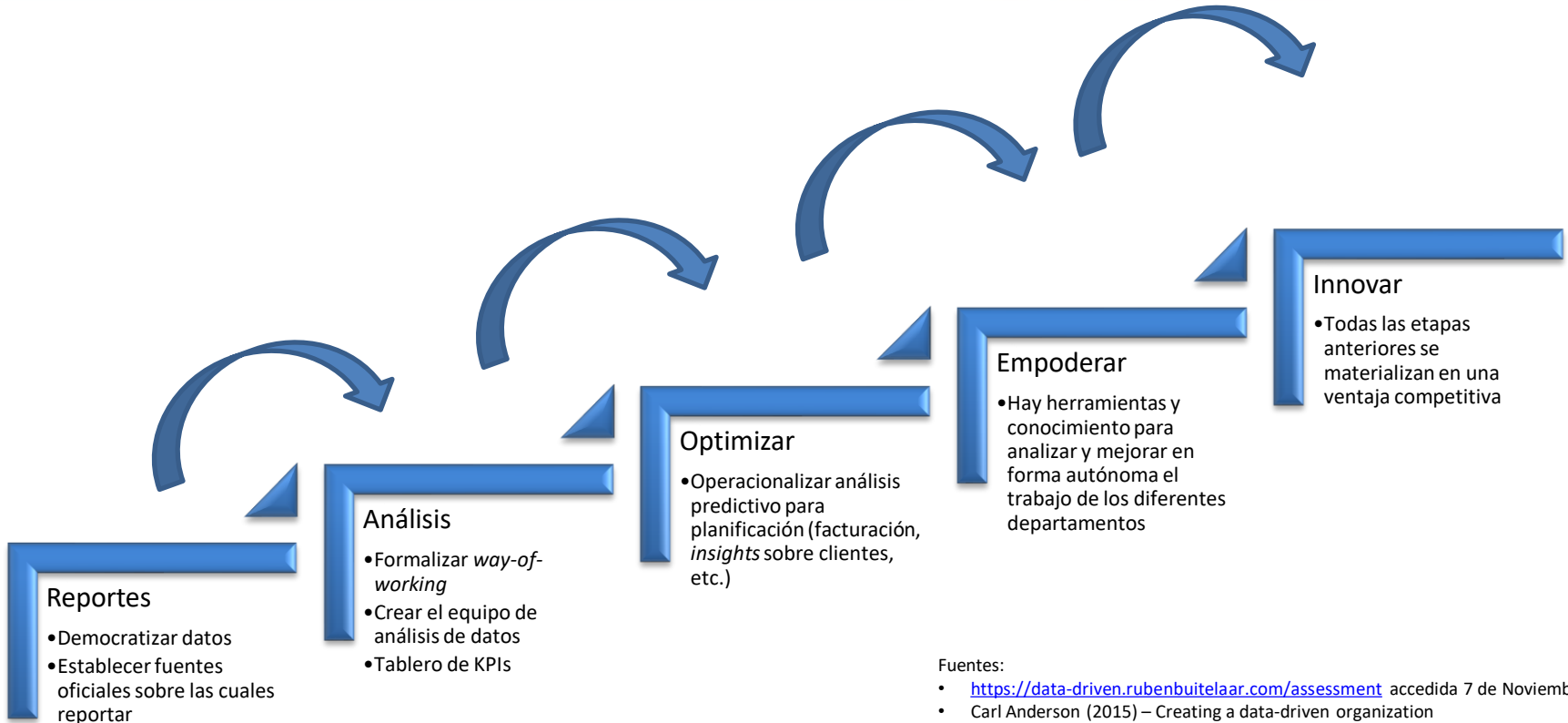
- Análisis de datos
 - Conceptos
 - Ejemplos
- Tiempo para trabajar en el caso

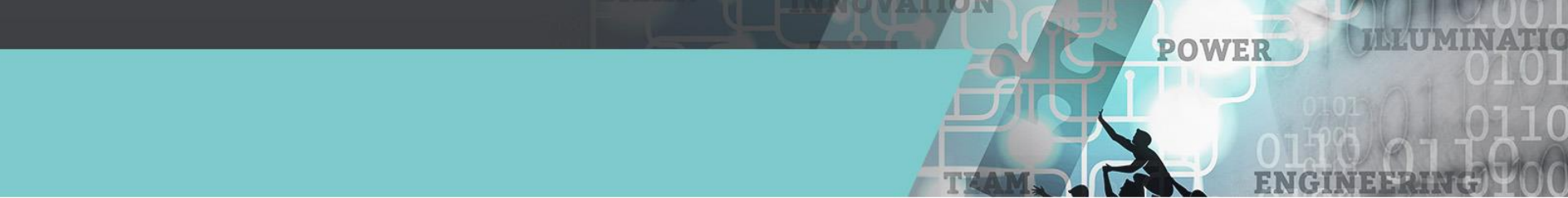
Cuatro dimensiones



- **Una visión y objetivos claros de negocios**
 - La visión debe resonar con los resultados del negocio
- **Uso de los datos**
 - Usar los datos para la toma de decisiones. Para crear la herramienta correcta es necesario promover la cultura de datos en primer lugar
 - La privacidad de los datos es importante. Las compañías que coleccionen y usan datos de sus clientes deben protegerlos
- **Motor de datos**
 - Evitar mirar la tecnologías con los lentes de la novelería
 - Armar un equipo fuerte con habilidades variadas (o un proveedor)
 - Los datos y su análisis no deben vivir en silos, deben de estar integrados en la operación de día a día
- **Ecosistemas de datos**
 - Es mas eficiente tomar datos de un ecosistema amplio para el intercambio y apoyo en el análisis

Modelo de madurez





Big Data e Inteligencia del negocio





Big Data

Unidades de almacenamiento

- En base binaria 2^n

Unidades de Medidas de Almacenamiento

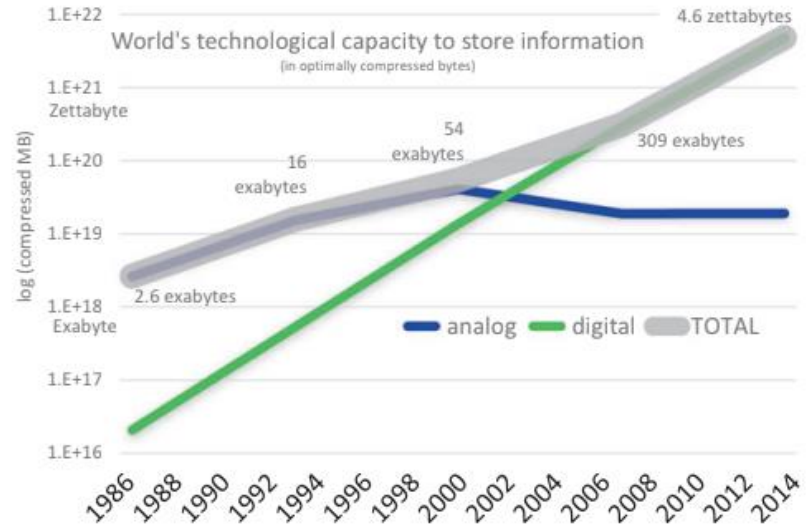
Medida	Simbología	Equivalencia	Equivalente en Bytes
byte	b	8 bits	1 byte
kilobyte	Kb	1024 bytes	1 024 bytes
megabyte	MB	1024 KB	1 048 576 bytes
gigabyte	GB	1024 MB	1 073 741 824 bytes
terabyte	TB	1024 GB	1 099 511 627 776 bytes
Petabyte	PB	1024 TB	1 125 899 906 842 624 bytes
Exabyte	EB	1024 PB	1 152 921 504 606 846 976 bytes
Zetabyte	ZB	1024 EB	1 180 591 620 717 411 303 424 bytes
Yottabyte	YB	1024 ZB	1 208 925 819 614 629 174 706 176 bytes
Brontobyte	BB	1024 YB	1 237 940 039 285 380 274 899 124 224 bytes
Geopbyte	GB	1024 BB	1 267 650 600 228 229 401 496 703 205 376 bytes

www.tiposdecomputadora.wordpress.com

Big Data

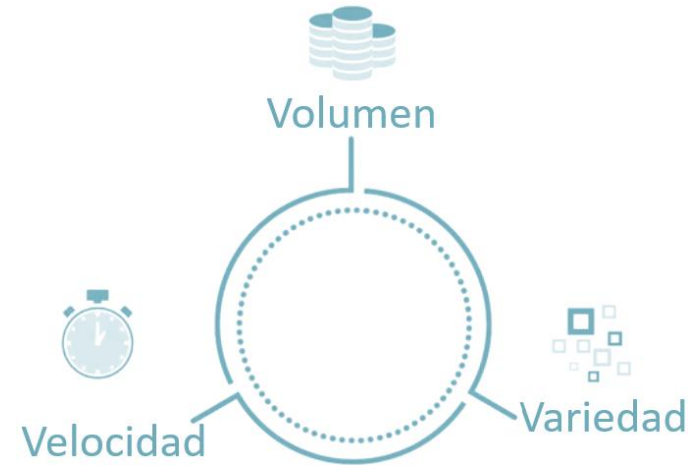
- La capacidad de almacenamiento digital ha aumentado en forma exponencial

Information Quantity,
Fig. 1 World's technological capacity to store information 1986–2014 (log on y-axis)
(Source: Based on the methodology of Hilbert and López (2011), with own estimates for 2007–2014)



Big Data - definición

- Se define mediante las tres “V”
 - Volumen.
 - Variedad.
 - Velocidad.
- Existen zettabytes de datos en formato no estructurado, que pueden intercambiarse en tiempo real, lo que contribuye al crecimiento de datos y por tanto de su volumen.



Big Data - aplicaciones



- Personalización de la oferta
 - Entender el cliente y ofrecer aquello que tiene mayor probabilidad de convertir
- Detección de fraudes
 - Detección de patrones de fraude por parte de empresas de pagos
- Mantenimiento preventivo
 - Prevenir problemas de estabilidad antes que ocurran

Big Data - desafíos

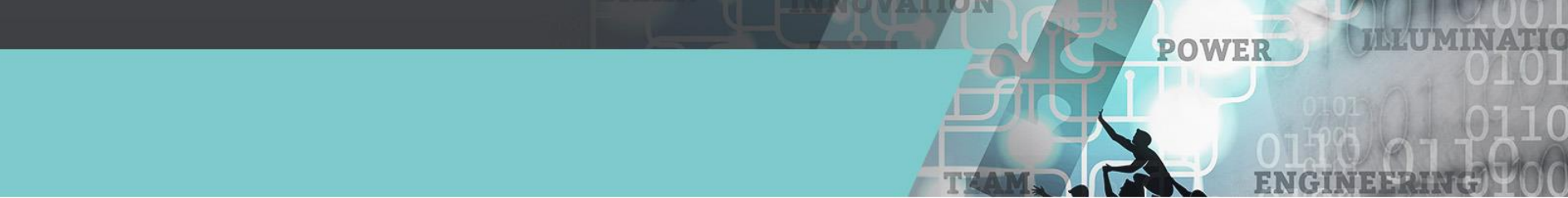
The background features a teal-to-white gradient with various technical and business-related terms in a light blue font, including 'INNOVATION', 'POWER', 'ILLUMINATION', 'TEAM', and 'ENGINEERING'. There are also binary code elements (0s and 1s) and a silhouette of a person climbing a ladder, symbolizing progress and data analysis.

- Usos no debidos de la información
- Confianza
- Privacidad

Big Data - resumen

The background features a teal-to-white gradient. On the right side, there is a silhouette of a person climbing a ladder towards a bright light. Various technical and business-related terms are scattered across the background, including 'INNOVATION', 'POWER', 'ILLUMINATION', 'TEAM', and 'ENGINEERING'. Binary code (0s and 1s) is also visible, particularly on the right side.

- La capacidad de almacenamiento crece exponencialmente
- Una cantidad enorme de datos no estructurados que se mueve a gran velocidad
- Se puede generar valor al trazar, conectar y analizar los datos
- No es solo cuestión de tecnología, sino de ganar la confianza de las personas



Inteligencia del negocio

Inteligencia del negocio



- A pesar del avance de la tecnología y las herramientas para asistir en el proceso de toma de decisiones, mucha gente toma decisiones por instinto.
- La inteligencia del negocio toma datos de ayer y hoy para tomar mejores decisiones sobre el mañana.



Decisiones a tiempo

- La inteligencia del negocio surge como un aliado en la vida de los negocios para la toma de decisiones.
- Ejemplos;
 - Un ejecutivo de ventas está deliberando entre que prospectos los ejecutivos de cuenta necesitan focalizarse en el último trimestre para aumentar las ganancias.
 - Un equipo de investigación y desarrollo está decidiendo que funcionalidades incluir en la próxima versión de un producto.
 - El departamento de fraude está decidiendo en cambios en el programa de fidelidad de los clientes que eliminara el riesgo de fraude sin sacrificar la satisfacción del cliente.

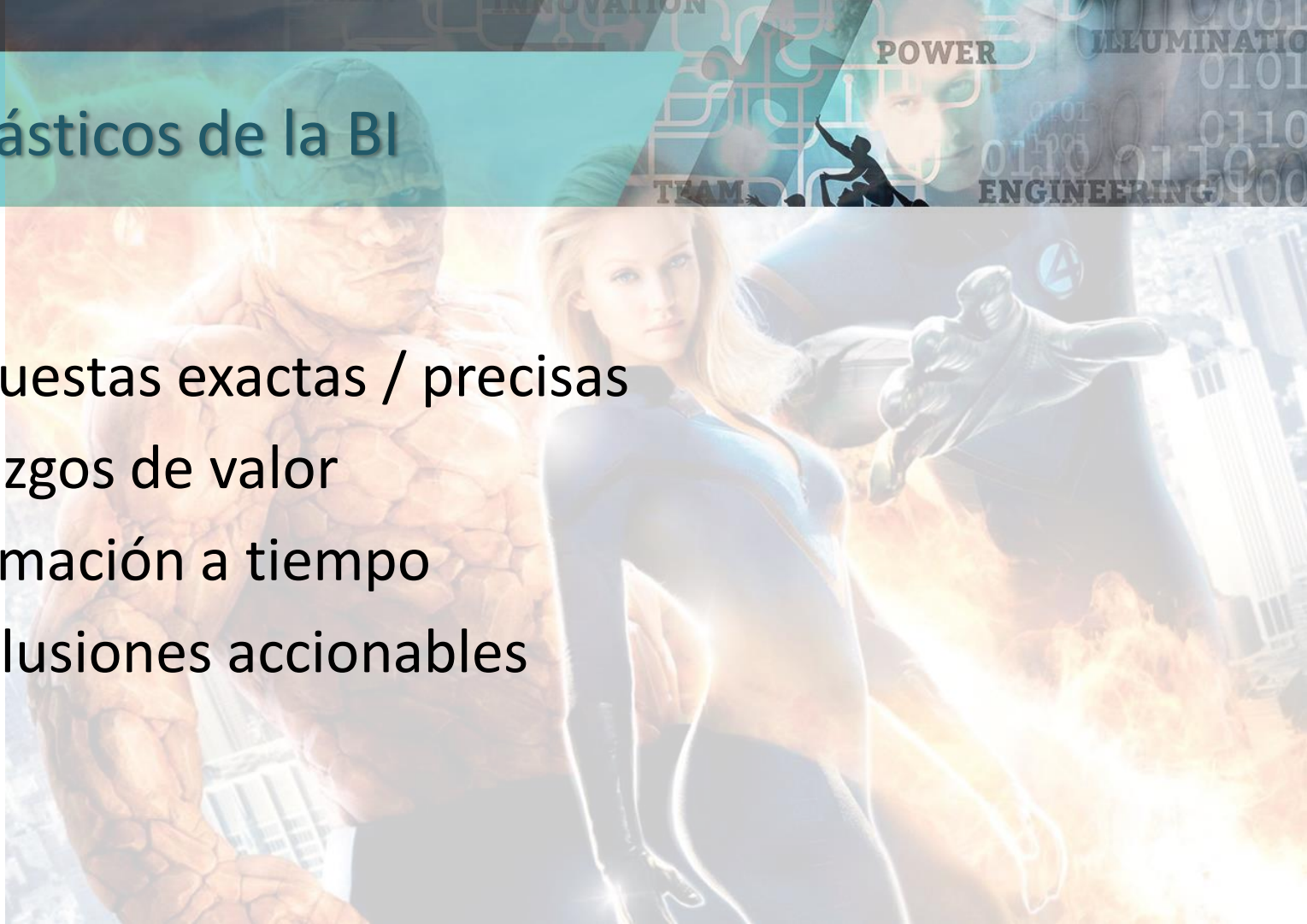
Inteligencia del negocio



- También conocida como BI, *bi-ai*.
- La inteligencia del negocio es esencialmente un conjunto de procesos y tecnologías que llevan a la obtención a tiempo de hallazgos de negocio que son accionables, exactos y de alto valor.
- Los hallazgos se les conoce también como *insights*

Los 4 fantásticos de la BI

1. Respuestas exactas / precisas
2. Hallazgos de valor
3. Información a tiempo
4. Conclusiones accionables



Los 4 fantásticos de la BI



1. Respuestas exactas / precisas.
 - Conclusiones derivadas de la realidad objetiva de la organización y de acuerdo a estándares rígidos de exactitud (por ejemplo; principios de gobernanza de la información).
2. Hallazgos de valor.
 - La información producida debe causar impacto material en la organización – aumento de facturación, reducción de costos u otro factor positivo.

Los 4 fantásticos de la BI

3. Información a tiempo.

- Los intervalos de tiempo en que se produce, computan y analiza la información deben ser lo suficientemente pequeños como para que los hallazgos sean relevantes, útiles y de valor para el o los decisores.

4. Conclusiones accionables.

- No solo alcanza que las conclusiones sean exactas, sino que tienen que ser un curso de acción ventajoso realizable. Debe ser posible moverse desde conclusión a acción.

Modelos de usuario de BI



- No es tan simple como ingresar información, apretar un botón y esperar por los resultados.
- Veamos algunos conceptos importantes

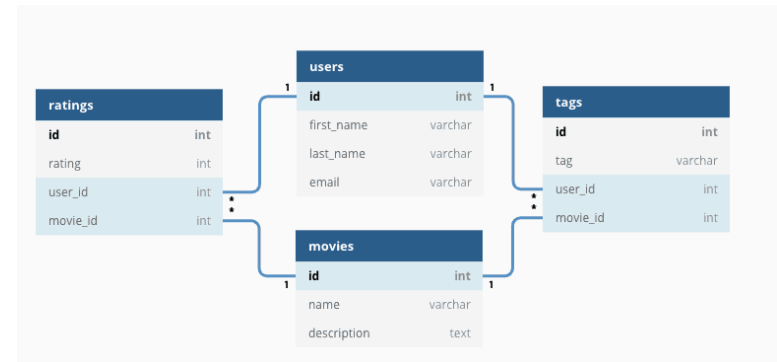


Algunos conceptos

- Una base de datos es una colección autodenominada de registros integrados. Es decir, contiene la descripción de su propia estructura como parte de los datos que almacena. Cuando decimos que los registros están integrados, decimos que existen relaciones entre los registros que los unen en una forma lógica y cohesiva.



Relaciones entre registros



Herramientas

- Herramientas de búsqueda (*query*): son diseñadas para permitirle al usuario crear una consulta para acceder a los metadatos (como los datos y tablas están relacionados).

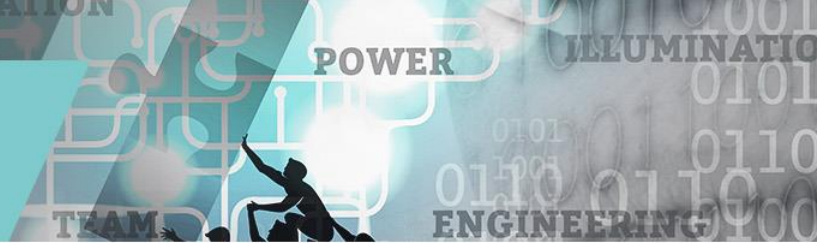
The screenshot shows the HeidiSQL interface with a table definition for 'Address' and its corresponding metadata table.

```
1 CREATE TABLE "Address" (  
2 "AddressID" INT NOT NULL,  
3 "AddressLine1" NVARCHAR(60) NOT NULL,  
4 "AddressLine2" NVARCHAR(60) NULL DEFAULT NULL,  
5 "City" NVARCHAR(30) NOT NULL,  
6 "StateProvinceID" INT NOT NULL,  
7 "PostalCode" NVARCHAR(15) NOT NULL,  
8 "SpatialLocation" NVARCHAR NULL,  
9 "rowguid" UNIQUEIDENTIFIER NOT NULL DEFAULT 'newid',  
10 "ModifiedDate" DATETIME NOT NULL DEFAULT 'getdate',  
11 PRIMARY KEY ("AddressID")
```

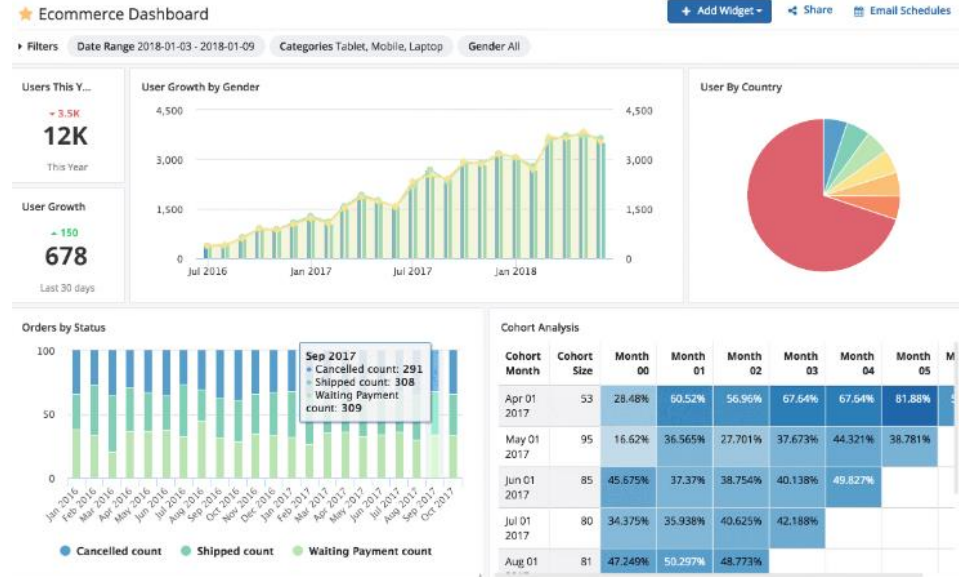
#	Name	Datatype	Length/Set	Unsign...	Allow N...	Zerofill	Default
1	AddressID	INT	10,0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	No default
2	AddressLine1	NVARCHAR	60	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	No default
3	AddressLine2	NVARCHAR	60	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	NULL
4	City	NVARCHAR	30	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	No default
5	StateProvinceID	INT	10,0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	No default
6	PostalCode	NVARCHAR	15	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	No default
7	SpatialLocation	NVARCHAR		<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	No default
8	rowguid	UNIQUEIDENTIFIER		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	newid
9	ModifiedDate	DATETIME	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	getdate

```
13 SELECT *, SCHEMA_NAME("schema_id") AS 'schema' FROM "AdventureWorks2012"."sys"."objects" WHERE "type" IN ('P', 'U', 'V', 'TR', 'FN'  
14 USE "AdventureWorks2012";  
15 SELECT * FROM INFORMATION_SCHEMA.COLUMNS WHERE TABLE_CATALOG="AdventureWorks2012" AND TABLE_NAME="Address";  
16 SELECT C.CONSTRAINT_NAME, C.CONSTRAINT_TYPE, K.COLUMN_NAME FROM INFORMATION_SCHEMA.TABLE_CONSTRAINTS AS C INNER JOIN INFORMATION_SC  
17 SELECT '' AS "Collation", '' AS "Charset", 0 AS "Id", '' AS "Default", '' AS "Compiled", 1 AS "Sortlen";
```

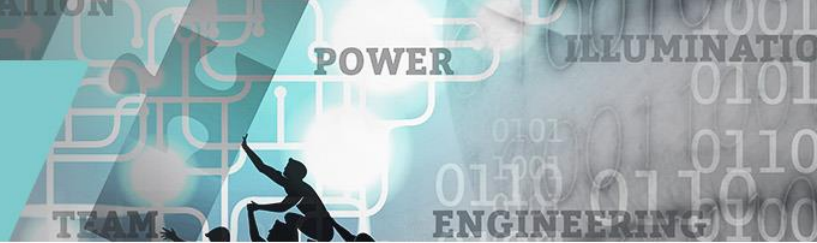
Herramientas



- Herramientas de reporte: ayudan a distribuir la información de las consultas en forma de reportes.



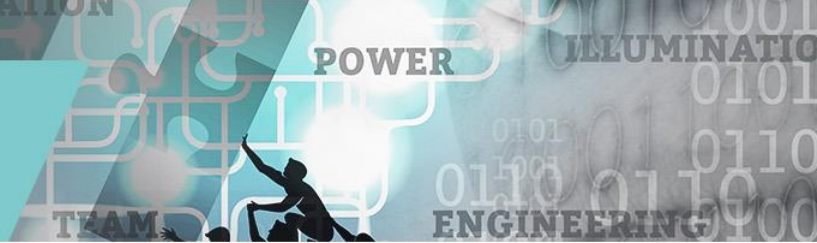
Datos unidimensionales



- Una tabla en bases de datos se refiere al tipo de modelado de datos donde se guardan los mismos
- Se dice que tiene una dimensión porque estamos mirando datos de ventas desde una sola perspectiva o dimensión; en este caso, por región.

<i>Annual Sales Data</i>	
<i>Region</i>	<i>Amount</i>
Northeast	\$45,091
Southeast	\$73,792
Central	\$88,122
West	\$63,054
TOTAL:	\$270,059

Dos dimensiones



- Pensar en las dimensiones como un sistema de coordenadas.

Table 5-4 Two Dimensions of Data Combined

Annual Sales Data

<i>Quarter</i>	<i>Product</i>		<i>Totals</i>
	<i>Gizmo</i>	<i>Widget</i>	
Q1	\$23,199	\$32,688	\$55,887
Q2	\$24,798	\$62,861	\$87,659
Q3	\$14,555	\$9,043	\$23,598
Q4	\$26,145	\$76,770	\$102,915
TOTALS	\$88,697	\$181,362	\$270,059

Tres dimensiones

Table 5-5 All Three Dimensions in One Table					
<i>Annual Sales Data</i>					
<i>Region</i>					
	<i>Northeast</i>	<i>Southeast</i>	<i>Central</i>	<i>West</i>	<i>Total</i>
GIZMOS					
Q1	\$1,543	\$14,098	\$1,991	\$5,567	\$23,199
Q2	\$6,811	\$2,822	\$13,300	\$1,865	\$24,798
Q3	\$5,190	\$5,050	\$2,106	\$2,209	\$14,555
Q4	\$2,347	\$8,005	\$5,900	\$9,893	\$26,145
WIDGETS					
Q1	\$3,555	\$5,520	\$6,828	\$16,785	\$32,688
Q2	\$9,158	\$15,999	\$18,096	\$19,608	\$62,861
Q3	\$2,486	\$1,297	\$4,247	\$1,013	\$9,043
Q4	\$14,001	\$21,001	\$35,654	\$6,114	\$76,770
TOTALS	\$45,091	\$73,792	\$88,122	\$63,054	\$270,059

Arquitectura BI

- *Online Analytical Processing (OLAP)*
- Una aplicación OLAP es un software diseñado para permitirle al usuario navegar, obtener y presentar datos de negocio
- Los cubos OLAP contienen datos multidimensionales
- Las herramientas de accos se usan para construir y preparar la información de forma adecuada para el análisis.

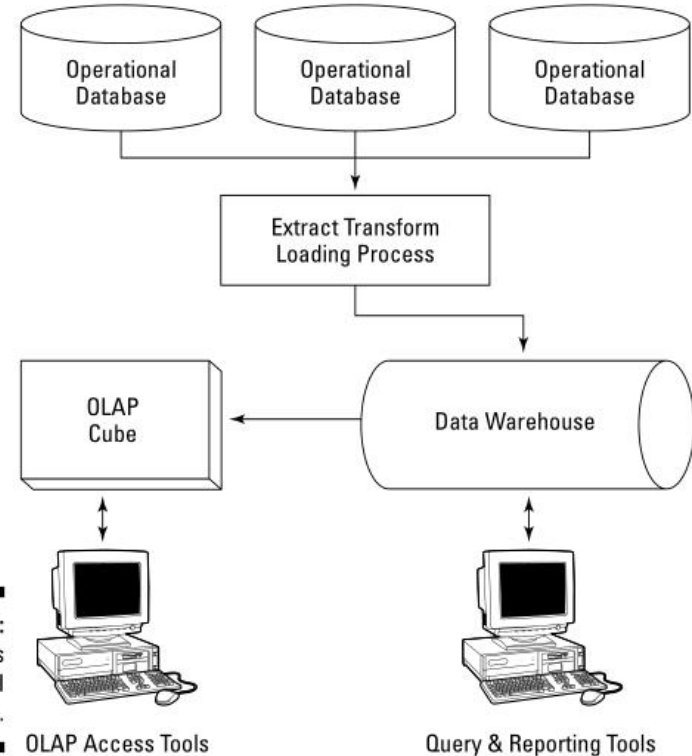
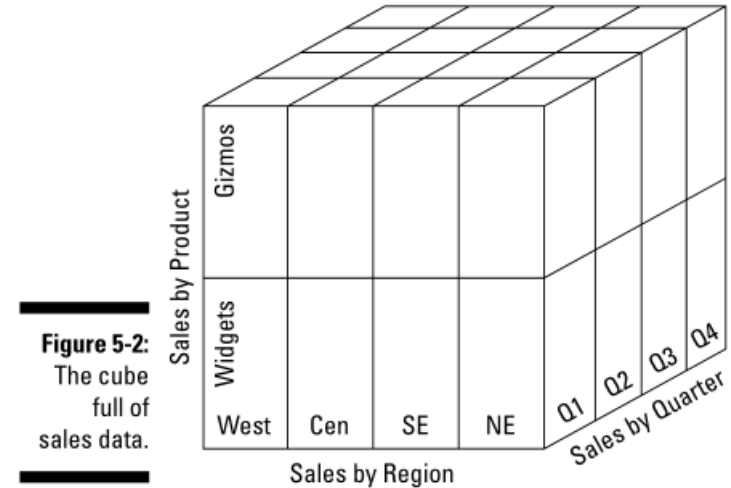


Figure 5-1:
OLAP tools
in the BI
architecture.

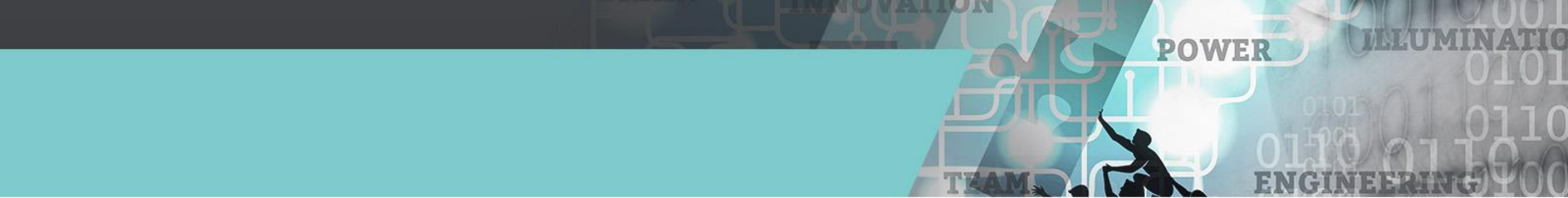
El cubo

- Para almacenar datos multidimensionales, se conceptualiza en un cubo, donde cada uno de los tres ejes representa una dimensión diferente de la información.
- Hay cubos de más de tres dimensiones



Utilidades de OLAP

- **Consolidación:** Recolectar datos en un nivel de abstracción superior.
- **Ir al detalle de los datos:** Poder abrir la información para obtener más detalles.
- **Computar:** La posibilidad de realizar operaciones matemáticas sobre los datos para simplificar.
- **Pivotear:** Ver los datos o un reporte desde diferentes perspectivas.



Cultura de datos

Cultura orientada a datos



- Una cultura orientada a datos es una cultura de confianza y apertura, y los datos están democratizados.
- Se provee acceso amplio a los datos a todos los individuos hasta donde es legalmente posible.
- Las organizaciones con una cultura orientada a datos tienen mayor potencial de impulsar la toma de decisión al menor nivel posible.
- Existe una amplia alfabetización en datos. Las personas son entrenadas en diseño, pensamiento crítico, presentación de los datos, uso de herramientas de BI, estadísticas, etc.

Cultura orientada a datos

- Orientación a metas, bajo una clara visión y dirección hacia adonde el negocio va.
- Cultura inquisitiva, se cambia la mentalidad de opiniones a hipótesis que puedan ser probadas objetivamente.
- Cultura de aprendizaje iterativo, hacer experimentos, instrumentar y medir resultados, interpretar los datos, realizar nuevas hipótesis y hacer nuevos experimentos.
- Una cultura anti-HiPPO (la opinión de la persona que recibe mejor paga), sin ningún reparo en los datos disponibles.

Modelo de madurez de cultura de datos



¿Qué pasará?

- Optimizar los procesos de negocios al traer hallazgos producidos por el análisis de datos en las operaciones

¿Por qué pasó?

- Profundizar en los datos para lograr hallazgos en el por qué las cosas ocurrieron

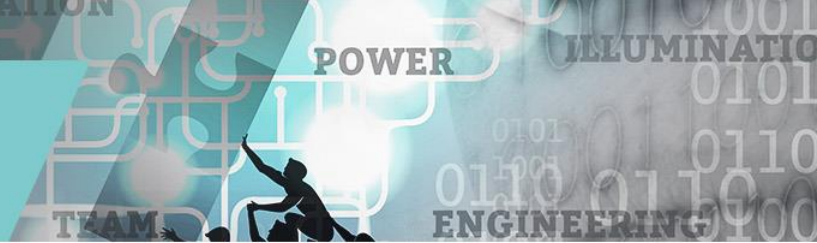
¿Qué pasó?

- Visualizar los datos existentes y crear las bases para un futuro 3DM

Ni Pálida Idea

- Las actividades de análisis de datos recién están iniciando
- Información es almacenada en silos y los reportes se realizan con mucho esfuerzo y de manera ad-hoc

Hoja de ruta de madurez



	M0	M1 Foco en ¿Por qué pasó?	M2 Foco en ¿Por qué pasó?	M3 Foco en ¿Qué pasará?
Personas		<ul style="list-style-type: none"> Adquirir talento externo o interno para construir un programa de análisis de datos Preparar y entusiasmar a los colaboradores por una cultura 3DM Remover preocupaciones por el cambio 	<ul style="list-style-type: none"> Entrene o atraiga científicos de datos para analizar con mayor profundidad los datos Retenga el talento con conocimiento crítico Introduzca a los colaboradores a las habilidades básica de datos Establecer una dirección clara desde arriba hacia abajo para alejarse de la vieja forma de trabaja Educar y proveer herramientas que permitan a dueños de procesos usar los datos y crear Dashboards 	<ul style="list-style-type: none"> Usar nuevas fuentes de Big Data en modelos de Machine Learning Usar modelos analíticos de aprendizaje para predecir el comportamiento de usuarios Estimular una cultura de mejora continua a través de los datos Motivar la experimentación para testear nuevas optimizaciones Proveer a los colaboradores con información en tiempo real sobre las operaciones y el desempeño Proveer a los colaboradores con la información para mejorar sus procesos
Procesos		<ul style="list-style-type: none"> Determinar datos y métricas críticas para operaciones y tomadores de decisión Estandarice reportes Defina responsabilidades básicas y dominios de decisión para evitar bloqueos y continuar evolucionando el cambio 	<ul style="list-style-type: none"> Recolectar y analizar datos relevantes Monitorear y asegurar la calidad de los datos Definir una estrategia para mantener los datos seguros y privados Definir métricas más avanzadas basadas en modelos analíticos de datos Crear un proceso para implementar recomendaciones basadas en evidencia a raíz del análisis de datos Crear reportes y Dashboards bajo demanda con la información correcta 	<ul style="list-style-type: none"> Recolectar datos de comportamiento de usuarios a gran escala Usar los datos de los clientes para personalizar la experiencia del cliente Definir y mostrar las métricas en tiempo real Automatizar el uso de hallazgos analíticos en los procesos existentes Enfocarse en transformar los datos automáticamente en acciones en vez de meros hallazgos
Tecnología		<ul style="list-style-type: none"> Identificar los requerimientos del negocio para armar una plataforma básica de análisis de datos Decidir si quiere construir o comprar software e infraestructura 	<ul style="list-style-type: none"> Profesionalizar la infraestructura de datos Estandarizar la tecnología, integraciones y Data Warehouse Adquirir tecnología que asiste a los científicos de datos analizar los datos 	<ul style="list-style-type: none"> Usar nuevas fuentes de Big Data en modelos de Machine Learning Usar modelos analíticos de aprendizaje para predecir el comportamiento de usuarios Usar el comportamiento predicho para optimizar la oferta