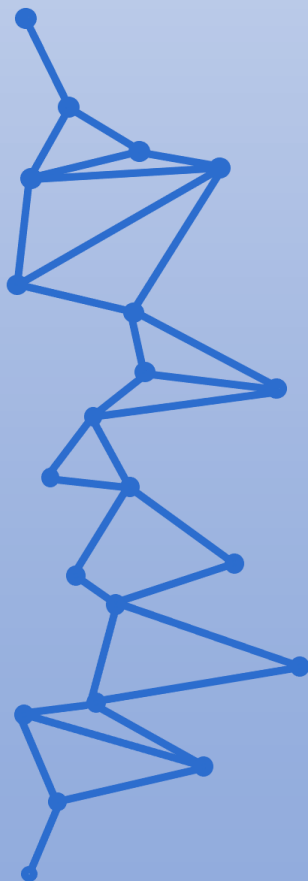




## Curso de Especialización de Inteligencia Artificial y Big Data (IABD)



### Sistemas de Big Data

UD05. BI, visualización de datos y cuadros de  
mando.  
Resumen.

JUAN ANTONIO GARCIA MUELAS

---

En **1958**, el investigador de IBM **Hans Peter Luhn publica** un artículo con el que se sientan las bases de la **Inteligencia de Negocio**.

La **Inteligencia de Negocio** (en inglés Business Intelligence o BI) comprende un **conjunto de tecnologías y técnicas** empleadas por las empresas **para realizar una correcta gestión** del negocio y una **toma inteligente de decisiones** basándose en los datos disponibles.

Se origina como **solución a problemas** comunes:

- ✓ **Respecto a la información.** Suele estar **diseminada en distintos sistemas y hojas de cálculo**, por lo que **no hay una única versión** de la verdad.
- ✓ **Respecto a la carga de trabajo de las personas.** Por los **cuellos de botella** creados al **solicitar informes o análisis** a las mismas personas.
- ✓ **Respecto al tiempo que tardan en realizarse las decisiones.** Por **esperar el informe** que llega tarde, porque quienes toman las decisiones **realizan sus propios informes retrasando la decisión**.
- ✓ **Respecto a la calidad de los informes.** Al **no haber una única versión**, los **resultados** pueden ser **dispares**.
- ✓ **Respecto al desperdicio de recursos.** Demasiados **informes que no llegan a usarse** o con **información repetida** que ya se conoce, o por ser realizados con sistemas operacionales que acaban ralentizando procesos vitales.

Una **implementación apropiada** de la Inteligencia de Negocio en una empresa produce **beneficios cuantificables**, que permiten tomar **mejores decisiones**, de un modo **más rápido y a un menor coste** (tanto económico como humano).

**Gracias a la Inteligencia de Negocio dejan de hacerse informes de modo manual y se puede mirar a más largo plazo para la toma de decisiones.**

Los **3 aspectos** primordiales o **requisitos de BI**:

- ✓ **Hechos:** son los **eventos** del negocio **acerca de los cuales queremos generar** informes o producir cualquier tipo de analítica. Ejemplos: Las ventas. Las compras. Los pagos. Los cobros. La facturación. El inventario de almacén. El precio del suelo. El valor de los activos de la empresa.
- ✓ **Medidas:** Son los **valores que se quieren analizar o reportar**. Algunas son relativamente sencillas de obtener mientras que para otras puede ser necesario realizar algún tipo de composición o procesamiento (por ejemplo, un cambio de divisa en función de sus valores relativos en cierto momento). Ejemplos: Los ingresos. Los costes. Los márgenes. Los beneficios. Días de pago.
- ✓ **Dimensiones:** Son las dimensiones del negocio **sobre las cuales se puede desear tomar porciones o filtrar**. Para cada una de esas dimensiones es necesario contar con los atributos correspondientes. Ejemplos: Por cliente. Por producto. Por fecha. Por área (país, ciudad, ...). Por vendedor. Por tienda. Por proveedor.

**Tipos de herramientas** para Inteligencia de Negocio:

- ✓ **Soluciones de Inteligencia de Negocio empresarial:** Son suites de Inteligencia de negocio con **gran cantidad de funcionalidades y muy escalables**, por lo que son capaces de analizar grandes cantidades de datos.
  - **Ventajas:** Muy escalables, por lo que pueden servir a los negocios más grandes. Tienen todas las características y funcionalidades que la empresa pueda necesitar.
  - **Inconvenientes:** Son **soluciones caras**, tanto de adquirir como de poner en marcha. Pueden resultar **complejas** para determinados usuario o negocios.

- ✓ **Inteligencia de Negocio incluida en ERP:** Son las **herramientas de BI incluidas en el ERP** que utiliza la empresa.
  - **Ventajas:** Están diseñadas para funcionar en el ERP al cual pertenecen, lo cual **facilita** gran parte del **trabajo**.
  - **Inconvenientes:** Por defecto **sólo muestran información** contenida en el ERP. Si pueden **extenderse** para poder adquirir **datos de otras fuentes** eso **suele ser complejo y caro** (típicamente pagando desarrollos a la medida por parte de consultores enviados por el vendedor del ERP).
- ✓ **Herramientas de Inteligencia de Negocio en la nube:** Son soluciones online que se emplean en modo software-como-servicio (**SaaS**), a las cuales se acceder de un modo seguro (encriptado) a través de Internet.
  - **Ventajas:** Su puesta en marcha supone un **menor riesgo económico**, ya que **se paga por uso**. No es necesario adquirir servidores propios ni tener que instalar software (al ser web). Pueden **accederse desde cualquier lugar con conexión** a Internet.
  - **Inconvenientes:** Algunas soluciones son demasiado **simples**. Siguen **requiriendo** cierto **trabajo** por parte de la empresa, como el referente a **Integración de Datos**. Al ser en la nube, **puede provocar dudas** a algunos usuarios **respecto a su seguridad**.

**Características en las herramientas para Inteligencia de Negocio:**

- ✓ **Seguridad.** Para que **solo** se pueda acceder a la **información de su rol o puesto**.
- ✓ **Capacidad para el autoservicio de informes.** Para **generarlos** de forma **sencilla o filtrar**.
- ✓ **Capacidad para hacer visualización de datos.** Para **representaciones gráficas**, facilitando la **toma de decisiones**.
- ✓ **Capacidad para el crear cuadros de mando.** Los cuadros de mando **permiten** que cada usuario pueda tener una **visión personalizada** de los **datos y la situación del negocio desde** el punto de vista de **su rol** en la empresa. **Son interactivos**, de modo que el usuario **puede realizar distintos filtrados** para poder ver la información que le interese en cada momento.
- ✓ **Planificación de informes.** Definir **informes** que serán enviados con cierta **periodicidad**.
- ✓ **Integración con Microsoft Office.** Que será la suite que normalmente use la empresa.
- ✓ **Soporte móvil.** Para poder **tomar decisiones** en cualquier **momento y lugar**.

**Datawarehouse** (almacén de datos), es un **repositorio** en el que una organización **guarda tanto información histórica como reciente**. Una instantánea denormalizada de la información almacenada en estructuras diseñadas para un **alto rendimiento en analítica**.

Un **datamart** contiene un **subconjunto de la información** que responde a las necesidades de un **departamento**, división o línea de negocio.

En ocasiones, existen en presencia de un datawarehouse (datamarts dependientes).

La opción de **datamarts independientes** tiene la **ventaja eliminar el datawarehouse central** y con ello disminuir en cierto modo la complejidad del sistema. Como **desventajas**, está que **no realiza informes o cuadro de mando transversales** (distintos departamentos), y deja de tener una **versión única**.

Un **datamart** es similar a un **datawarehouse** pero **con información de un departamento**, división o línea de negocio.

Cuando una empresa tiene un sistema con datamarts independientes significa que **no se alimentan desde un datawarehouse corporativo**.

Un gráfico de barras apiladas puede mostrarse en horizontal o en vertical.

Los **mapas de calor** tienen utilidad cuando los **datos** de entrada tienen **3 dimensiones**.

Cuando hablamos de gráficos circulares dibujar distintas series dentro de un mismo diagrama no suele dar muy buenos resultados visuales.

Los gráficos circulares son muy útiles cuando queremos discernir qué partes de un total caen dentro de diversas categorías.

El gráfico de burbujas es similar al de puntos, con la diferencia de que los puntos (en este caso burbujas) además tienen distintos tamaños.

Los gráficos de cajas muestran datos atípicos, máximos, mínimos y medianas.

Cuando hablamos de visualización de datos, los indicadores se utilizan para indicar medidas relevantes, en ocasiones dentro de rangos.

#### Diferencias entre informes y cuadros de mando

Informes:	Cuadros de mando:
<ul style="list-style-type: none"><li>✓ Más enfocados a instantánea que a análisis en tiempo real.</li><li>✓ Más enfocados a página.</li><li>✓ Más enfocados a tablas.</li><li>✓ Ocupan más espacio (quizás varias páginas).</li><li>✓ Con mayor cantidad de datos históricos.</li><li>✓ Suelen enviarse de forma periódica.</li></ul>	<ul style="list-style-type: none"><li>✓ Dinámicos: muestran cambios en tiempo real.</li><li>✓ Más interactivos.</li><li>✓ Más enfocados a pantalla.</li><li>✓ Más enfocados a gráficos de visualización.</li><li>✓ Más resumidos (a ser posible caben en una pantalla).</li><li>✓ Intentan mostrar lo que está ocurriendo con un rápido vistazo.</li><li>✓ Pueden mostrar alarmas cuando cierta métrica cambia o llega a determinado valor.</li></ul>

Los **cuadros de mando**, de modo general son **más interactivos que los informes**.

Los **informes**, de modo general están **más enfocados a instantánea que a análisis en tiempo real**.

**Factores** que considerar al **crear un cuadro de mandos**:

- ✓ **Con qué tipo de datos se está trabajando:** El tipo de datos con los que trabajamos influye en el modo en el que son mejor presentados y visualizados.  
Lo más común es **trabajar con hechos**, y en tal caso la mayoría de la **información** a mostrar es **cuantitativa**. Cuando no es el caso, debe tomarse en **consideración cómo** vamos a **mostrar** tal **información no cuantitativa junto con la que sí lo es**.
- ✓ **Cómo de actuales deben ser los datos:** La **frecuencia con la que** el cuadro de mando **necesita ser actualizado** puede tener un efecto significativo en el modo en el que lo diseñaremos.  
En ocasiones es necesario mostrar información en cierta medida **desfasada** para tener tiempo de realizar todo el procesado. Si es así, el **usuario debe ser consciente de cómo de antiguos son** los datos que está viendo.  
Si el usuario necesita tener los datos actualizados prácticamente **en tiempo real**, eso impone una **limitación a la riqueza de la información** que se puede mostrar **y a la cantidad de datos en la que se basa**.

- ✓ **Cuántos usuarios (o qué tipo de usuarios) lo van a usar:** Un cuadro de mando es **más efectivo cuanto más personalizado está** para cada posible usuario. Sin embargo puede haber **demasiados usuarios**, en cuyo caso puede ser interesante **crear tipologías y simplemente crear un cuadro de mando para cada tipología**.  
En todo caso, es importante que los cuadros de mando sean lo suficientemente flexibles para que cada usuario pueda terminar obteniendo aquella información que desea.
- ✓ **Cómo de experimentados son los usuarios:** Es **clave a la hora de determinar cómo de complejo o detallado será** el cuadro de mando a crear.  
De cualquier modo, siempre es **recomendable** que el cuadro de mando sea lo suficientemente **sencillo** como para que lo pueda utilizar el usuario menos experimentado, pero que permita profundizar y filtrar al usuario más experto.
- ✓ **Cómo se accederá al cuadro de mando:** Es importante **saber de antemano** desde qué **plataforma/s** se accederá al cuadro de mando, ya que el diseño debe ser enfocado de un modo específico si habrá usuarios que lo accederán desde terminales móviles (con pantallas más pequeñas).

A la hora de **definir qué incluiremos** en nuestro cuadro de mando:

- ✓ **Definir qué quiere conseguir el cuadro de mando:** En la práctica en las empresas se construyen muchos cuadros de mando sin hacerse antes la pregunta de qué se quiere conseguir exactamente con cada uno de ellos.  
En concreto debemos compilar la lista de requisitos para saber qué información es la que el usuario quiere conocer basada en los datos.
- ✓ **Evitar concentrarse en la apariencia:** El estilo y la apariencia son importantes para obtener una experiencia gratificante, pero hay que tener en cuenta que el fin del cuadro de mando es mostrar la información relevante que el usuario necesita, lo cual no puede dejarse en ningún momento de lado a cambio de una apariencia vistosa.  
Al crear un cuadro de mando lo más importante es **mostrar la información relevante**, a ser posible **con buen estilo y apariencia**.
- ✓ **Determinar qué información es realmente necesaria:** Dado que el cuadro de mando ideal ocupa poco espacio para poder comprobarlo de un vistazo, es importante determinar qué cosas realmente necesitan estar en él (algo que por lo general no ocurre en un informe, ya que típicamente ocupan más espacio).  
Será el usuario final quien decida qué necesita ver, aunque quien crea el cuadro de mando puede guiarlo para tomar la mejor decisión a este respecto.

**Métricas** a emplear cuando se crea un cuadro de mando:

- ✓ Finanzas: Ingresos, gastos, beneficios.
- ✓ Ventas: facturación, reservas, pedidos, precios de venta...
- ✓ Soporte técnico: Número de llamadas al soporte, casos resueltos, duración de llamadas...
- ✓ Marketing: Cuota de mercado, éxito de campaña, demografía compradores...
- ✓ Cumplimiento: retrasos, niveles de inventario...
- ✓ Fabricación: Unidades fabricadas, tiempos de fabricación, defectos...
- ✓ Recursos humanos: satisfacción empleados, rotación, posiciones abiertas...

**Características** deseables en una **métrica**:

- ✓ **Accionable:** Dado que el fin de la Inteligencia de Negocio no es meramente informativo sino sobre todo el de ser capaces de tomar buenas decisiones, es importante mostrar métricas ante las cuales sepamos cómo actuar si queremos modificar su tendencia. Lo opuesto son

las llamadas 'métricas de la vanidad', las cuales simplemente muestran tendencias en los datos ante los cuales no se sabe qué hacer.

Que una **métrica** sea **accionable** significa que sepamos **cómo actuar** si queremos **modificar** su **tendencia**.

- ✓ **Transparente:** Implica **que se sepa** a ciencia cierta **de dónde proviene la información** en la que se basa. Para ello es importante que las métricas se produzcan mediante **cálculos** lo más **simples** posible, de modo que el usuario pueda seguirlos.
- ✓ **Accesible:** Implica que la métrica se produzca a partir de **datos sencillos de obtener y de mantener**. Si cada vez que es necesario obtener tales datos hay que realizar un proceso costoso, ello va en detrimento de la frecuencia a la cual puede actualizarse el cuadro de mando. Ello reduce su efectividad y su relevancia, al obligar a trabajar con datos con mayor desfase.
- ✓ **Reconocible:** Implica **que no exista ambigüedad sobre qué significa**. Si es subjetiva o difícil de comprender ello va en detrimento de su correcta interpretación.

Trabajando con Power BI, si queremos relacionar dos campos de distintas tablas Power BI creará automáticamente la relación si tienen el mismo nombre, y si no lo hacemos manualmente en la vista "Modelo".

Cuando en un cuadro de mando de Power BI hacemos clic sobre una porción en un gráfico circular toda la información se filtra de modo que sólo vemos la referente al significado de la porción seleccionada.