**Datos de identificación del programa de formación**

| PROGRAMA DE FORMACIÓN | Gestión de analítica de datos |
| --- | --- |

| COMPETENCIA | 220501114 - Sistematizar datos masivos de acuerdo con métodos de analítica y herramientas tecnológicas. | RESULTADOS DE APRENDIZAJE | 220501114-01. Preparar los datos, conforme a metodologías de analítica de datos. |
| --- | --- | --- | --- |

| NÚMERO DEL COMPONENTE FORMATIVO | CF10 |
| --- | --- |
| NOMBRE DEL COMPONENTE FORMATIVO | Análisis y conocimiento de datos enfocado a la analítica. |
| BREVE DESCRIPCIÓN | El recurso educativo presenta una introducción amplia y técnica sobre conceptos, teorías y herramientas empleadas en sistematización de datos. Se dan las pautas de las metodologías y paradigmas usados para la gestión de información, y la transformación y carga que deben realizarse con las tecnologías disponibles. |
| PALABRAS CLAVE | Analítica, datos, inteligencia de negocio. |

| ÁREA OCUPACIONAL | Procesamiento, fabricación y ensamble. |
| --- | --- |
| IDIOMA | Español. |

**Tabla de contenidos**

[**Introducción 4**](#_heading=h.3j2qqm3)

[**1**](#_heading=h.1y810tw) **Los datos 5**

[1.1 Tipos de datos 8](#_heading=h.4i7ojhp)

[1.2 Ecosistemas de datos 16](#_heading=h.2xcytpi)

[**2**](#_heading=h.1ci93xb) **Calidad de los datos (*Quality data*) 23**

[2.1 Dimensiones de la calidad de los datos 26](#_heading=h.3whwml4)

[2.2 Datos maestros (](#_heading=h.2bn6wsx)*Master Data*) 30

[**3**](#_heading=h.qsh70q) **Arquitectura analítica de datos 32**

[3.1 Conceptos básicos estadísticos 33](#_heading=h.3as4poj)

[3.2 Modelos matemáticos 44](#_heading=h.2p2csry)

[3.3 Gestión de datos 49](#_heading=h.147n2zr)

[**4**](#_heading=h.3o7alnk) **Generalidades: Herramienta inteligencia de negocio 53**

[4.1 Instalación 53](#_heading=h.23ckvvd)

[4.2 Reportes (](#_heading=h.ihv636)*Data Source*) 54

[4.3 Desarrollo y representación gráfica de datos 58](#_heading=h.1hmsyys)

# Introducción

| Cuadro de texto |
| --- |
| Estimado aprendiz, reciba una cordial bienvenida a este recurso de aprendizaje orientado al análisis y conocimiento de datos enfocado a la analítica. Durante su desarrollo se abordarán los ejes temáticos datos, calidad de los datos, arquitectura de los datos y herramientas para la inteligencia de negocios. Al finalizar estará en capacidad de implementar pequeños proyectos de inteligencia de negocios conforme a los requerimientos. Antes de empezar, se sugiere revisar el siguiente video para identificar el contexto de aprendizaje:  ¡Le deseamos una experiencia de aprendizaje significativa y memorable! |

**Ver Anexo1: CF10-Guion video introductorio**

**Desarrollo de contenido**

# Los datos

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| Cuando el ser humano empezó a diferenciarse de las demás especies, una de sus características más importantes fue y es, la manera de comportarse socialmente; es decir, cuando en las relaciones sociales y comportamentales entre manadas los integrantes empezaron a cuidarse, a comunicarse, etc. Desde ese momento la información y, la manera en que la misma configura los datos a través del proceso comunicativo, generó las diversas maneras de comunicarse e incluso miles de mensajes de las primeras civilizaciones han quedado guardados por miles de años hasta nuestros días a través de pinturas, esculturas y símbolos.    Imagen sugerida de primeros datos (arte rupestre) en: <https://www.shutterstock.com/es/image-photo/archaeological-prehistoric-human-cliff-paint-over-100565722>  Imagen: 228131\_i94 | |

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| Bajo este contexto es importante recordar la diferencia entre datos e información.  Para entender esto se puede partir del ejemplo de que cuando las civilizaciones descubren mensajes de humanos antiguos no tienen ninguna forma de interpretar los mensajes; ante ojos no entrenados lo encontrado no son más más que símbolos con valor artístico y a pesar de tener la cognición antropológica de que los mensajes representan cosas, estos realmente serían solo datos (símbolos) sin información (no interpretados).  Imagen: 228131\_i95  Imagen sugerida de primeros datos (escritura cuneiforme) en: | |

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| En conclusión, los datos no son un asunto nuevo, sino que surgen a partir de la misma evolución humana, en donde se van volviendo complejos y bastos conforme avanzan las civilizaciones. Es por esto por lo que en el ámbito digital y tecnológico actual se convierten en un asunto de estudio profundo y técnico, pues el desarrollo humano y productivo tiene como plataforma principal el uso de los datos que constituyen información; así mismo la información se convierte en conocimiento, el cual es el componente principal para tomar acciones, las cuales a su vez generan nueva información.  Imagen sugerida de evolución de la información y datos. Construcción propia  Imagen: 228131\_i96 | |

## Tipos de datos

| **Tipo de recurso** | Slider pasos | |
| --- | --- | --- |
| **Introducción** | Los datos se pueden clasificar según su formato o el tipo de información que contienen, lo que sirve específicamente para la estructuración de arreglos de los mismos o la creación de bases de datos.  En este sentido, cada lenguaje de programación y motor de base de datos puede tener sus propios tipos de datos; sus características se muestran a continuación de manera general, resaltando solo las más comunes. | |
| **Slide 1** | En un nivel inferior de máquina los compiladores o procesamientos de datos reservan una cantidad de *bits,* según el tipo de datos que se seleccionen; por ello no solo es importante identificar la naturaleza de los datos a almacenar ya sean números, fechas, texto, etc. sino también determinar qué tantos recursos emplea cada tipo de datos desde el punto de vista de memoria, pues de esto también depende la velocidad y agilidad de procesamiento. | Imagen: 228131\_i97  Imagen que representa manejo informático de diferentes tipos de datos  <https://www.shutterstock.com/es/image-vector/data-analysis-database-visualization-young-woman-1995137660> |
| **Slide 2** | **Nota: se puede incrustar aquí el recurso de Carrusel de Tarjetas elaborado a continuación de este recurso.**  Si no se puede incrustar, entonces construir un gráfico animado con esta información: | |
| **Figura 2**  *Tipos de datos*    **Imagen: Tipos de datos, construcción propia**  Imagen: 228131\_i98 | |
| **Slide 3** | **Tabla tipos de datos**  **Tabla 1** *Tipos de datos*   | Tipo de dato | Denominación común | Tamaño en *BITs* | Rango de valores o capacidad | | --- | --- | --- | --- | | Enteros: números sin decimal, estos pueden ser positivos o negativos. Existen varios subtipos | *BIT* | 1 *bit* | 0, 1 | | *TINYINT* | 1 *byte* | 0 a 255 | | *SMALLINT* | 2 *bytes* | (-32,768 a +32,768) | | *INT* | 4 *bytes* | -231 hasta 231 | | Real | *REAL* | 4 *bytes* | -3.40E+38 hasta 3.40E+38 | | *FLOAT* | 4 a 8 *bytes* | -1,79E+308 hasta 1,79E+308 | | *MONEY* | 8 *bytes* | 922.337.203.685.477,5808 hasta 922.337.203.685.477,5808 | | Cadena de texto | *STRING* | 2 *bytes* | 231 caracteres Unicode. | | Tiempo | *TIME* | 5 *byte* | 00:00:00.0000000 hasta 23:59:59.9999999 | | DATE | 3 *bytes* | 0001-01-01 hasta 31.12.99 | | *SMALLDATETIME* | 4 *bytes* | 1900-01-01 hasta 2079-06-06 | | *Date time* | 8 *bytes* | 1753-01-01 hasta 9999-12-31 |   *228131\_T1. Adaptada de tipo de d* | |
| **Slide 4** | Si bien este componente formativo no está orientado al diseño de base de datos o programación, tener claridad acerca de los tipos de datos, sus rangos y usos es fundamental al momento de planificar la gestión de datos y la modelación de la información orientada a la analítica. | Imagen: 228131\_i99  Se puede editar imagen con frase de “Mensaje importante” en español. |
| **Slide 5** | En síntesis, la analítica de datos responde a preguntas de hechos del negocio, tales como;   * ¿Qué se hizo (filtros)? * ¿Cuánto se hizo (operaciones matemáticas con campos numéricos)? * ¿Cuándo (fechas y tiempos)?   En este orden los cálculos, filtros y medidas se realizan según el tipo de datos.  Posteriormente se detallarán algunas de estas operaciones enfocadas a la analítica | Imagen: 228131\_i100 |

## Ecosistemas de datos

| **Tipo de recurso** | Acordeón tipo 2 |
| --- | --- |
| **Introducción** | Se refiere a un conjunto de infraestructura de sistemas de información de varias áreas de una organización; en este ecosistema se incluyen todas y cada una de las diferentes interacciones que se dan, tanto entre los colaboradores y procesos como entre estos y el entorno en que se encuentran (clientes proveedores, sedes, etc.) |
| **Figura 3** *Ecosistema de actores de datos abiertos*    Imagen: 228131\_i106  **Imagen de guía: ejemplo de Ecosistema de datos, tomado de:**  [**https://herramientas.datos.gov.co/sites/default/files/Guia%20de%20Datos%20Abiertos%20de%20Colombia.pdf**](https://herramientas.datos.gov.co/sites/default/files/Guia%20de%20Datos%20Abiertos%20de%20Colombia.pdf) **pág 14** | |
| **Implican.**  Los ecosistemas recogen interacciones entre procesos y colaboradores, lo que implica tener datos y políticas tecnológicas integradas. Incluso toma elementos e interacciones externas para ampliarse a los ecosistemas compartidos para mejorar la productividad y conexión con otras fuentes pertinentes. | |
| **Cubren.**  En otras palabras, no solo abarca la integración de aplicaciones sino lo más importante, la integración de la organización; existe una única opinión en relación con que las empresas necesitan realizar prácticas de integración para tener éxito. Sin embargo, pocas empresas logran con eficacia obtener estas integraciones. | |
| **Utilidad.**  Los ecosistemas de datos o ecosistemas digitales conectados son necesarios para construir conexiones, impulsar la transformación y mejorar los resultados. Una empresa que tiene todos sus procesos funcionando de manera articulada es aquella que aprovechará todo lo que la transformación digital tiene para ofrecer. (RSM, 2020) | |

**Ventajas de los ecosistemas de datos**

| **Tipo de recurso** | | Pestañas o *tabs* Verticales |
| --- | --- | --- |
| **Introducción** | | Un ecosistema de datos permite a las organizaciones compartir datos y conocimientos relevantes dentro de las regulaciones locales aplicables, creando un nuevo valor para todos los participantes. Esto brinda beneficios a las empresas que conducen a ahorros de costos, nuevas fuentes de ingresos y mayor productividad. A continuación, conoce más beneficios de estos ecosistemas. |
| **Imagen: Imagen que representa “ventajas”:**  [**https://www.shutterstock.com/es/image-photo/businessman-man-hold-hand-offer-positive-19**](https://www.shutterstock.com/es/image-photo/businessman-man-hold-hand-offer-positive-1925310017) | | |
| **Mejoramiento de los procesos** | **Mayores ingresos al mejorar;**   |  | La agilidad en los procesos ahorrándose el tiempo y trabajo de registrar los datos más de una vez. | | --- | --- | |  | La experiencia de usuarios, pues en sistemas ya no se trabajará por área, sino como empresa integrada. En este sentido, cuando el cliente se acerque a realizar solicitudes o compras, los datos básicos se tendrán, independiente del área; esto representa más dinero y la posibilidad de atender más clientes con los mismos o menos recursos. | |  | El flujo de trabajo empleado, medios y datos digitales, que ayuda a suprimir errores ahorrando tiempo y haciendo las tareas de manera más eficaz y eficiente, lo que se traduce en ahorro de dinero. | |  | Los canales de comunicación y las relaciones laborales con los clientes y socios y reduce los costos operativos debido a los procesos de datos automatizados y la eficiencia en toda la empresa. | | |
| **Crea nuevas oportunidades de negocio** | **Servicio de valor agregado.**  Cuando la información está al alcance y hay una interpretación juiciosa se pueden observar y analizar los datos completos que fluyen a través de la organización y aprovecharlos para crear nuevos productos o servicios.  Los ecosistemas de datos y ecosistemas tecnológicos no solo mejoran los procesos actuales de generación de ingresos, sino que también crean servicios de valor agregado para nuevas oportunidades de ingresos. | |
| **Aumentan la velocidad de innovación y tecnología** | **Adopción de nuevas tecnologías.**  Una estrategia de integración de ecosistemas permite a las organizaciones adoptar completamente las nuevas tecnologías de manera que se minimicen los procesos de curvas de aprendizaje y se materialice la verificación y contrastación con el mercado de manera ágil; además permite tomar decisiones con mayor agilidad. (RSM, 2020). | |

**Aspectos a tener en cuenta en los ecosistemas de datos**

| **Tipo de recurso** | Rutas / Pasos. Verticales 1 |
| --- | --- |
| **Introducción** | En general la implementación de una estrategia de transformación digital implica tener ecosistemas digitales y de datos y además requiere de varios aspectos a tener en cuenta: |
| Imagen: 228131\_i108  **Imagen sugerida:** representación de ecosistemas digitales (se puede recrear de otra forma). Fuente: <https://www.pragma.com.co/blog/que-es-un-ecosistema-digital> | |
| **Botón 1** | **La naturaleza del negocio**  Ocasionalmente es importante definir si la organización hace parte directa de la industria tecnológica o se dedica a producir o comercializar productos físicos o de transformación; en caso de que sea una empresa manufacturera, por ejemplo, se debe diseñar la manera en que cada proceso pueda generar datos. En este sentido, lo ideal es la implementación de un software, sin embargo, no siempre es posible, así que si se modelan procesos donde los hechos se puedan registrar en hojas de Excel u otros recursos informáticos disponibles estaría bien; lo importante es garantizar estructura y calidad de la información. |
| **Botón 2** | **Tamaño de la empresa**  No es igual el proceso para grandes corporaciones que por lo general ya tienen una experiencia y avance importante en su transformación digital y cuentan con un ecosistema digital y de datos muy avanzado.  Lo importante es definir la necesidad para saber si lo que se requiere es modelar procesos en áreas que aún no se integran al ecosistema o definir qué otros procesos de reportes faltan, por ejemplo, los que se deben generar para aprovechar mejor el flujo de datos disponibles en las bodegas de datos. |
| **Botón 3** | **Recursos económicos disponibles**  Si bien la transformación digital implica una ganancia de dinero, esta curva no se empieza a ver de manera inmediata. Es importante tener mediciones claras de indicadores concretos y recursos invertidos a lo largo del tiempo; solo de esta manera es posible visualizar el ahorro y ganancias que las empresas experimentan. En este sentido, aunque la inversión inicial es alta luego se compensa con la línea de costos hasta sobrepasarla y generar más ganancias. |
| **Botón 4** | **Conocimiento técnico y apoyo en estrategias de transformación digital**  En el caso de que no sea una empresa relacionada con las tecnologías, no es rentable contratar un equipo de trabajo en tecnología, la mejor alternativa es contratar una empresa que realice estos desarrollos en modalidad *outsourcing*, sin embargo, es muy importante que desde la empresa se cuente con un experto en tecnología y que represente los intereses de la organización; este profesional, debe describir clara y técnicamente los requisitos y el apoyo de desarrollo que demandan las necesidades de la empresa. |
| **Botón 5** | **Definir infraestructura digital actual y acceder a los datos que se generan**  Un aspecto muy importante, es evaluar y entender la infraestructura tecnológica que se ha venido usando, se deben realizar las gestiones administrativas y técnicas para tener acceso a los datos generados. Si es un proveedor de *software*, realizar las gestiones ante terceros, como revisión de contratos y normatividad vigente para exigirle a estos que faciliten los datos de la organización. |
| **Botón 6** | **Definir arquitectura local, nube o híbrida**  Este es un aspecto importante ya que en él se define la arquitectura tecnológica y de esto dependen los recursos requeridos para su desarrollo.  La tendencia actual para la gestión de datos e infraestructura tecnológica es emplear servicios en la nube y por lo general los proveedores actuales tienen muy buenas capacidades y presentan confiabilidad en los aspectos de seguridad, disponibilidad y confidencialidad. Sin embargo, este análisis depende de cada empresa; ocasionalmente cuando ya tienen sistemas de información propios lo mejor es mantenerlos de esa manera, pero tener siempre en mente hacia el futuro hacer la migración a arquitecturas de nube. |
| **Botón 7** | **Definir proveedor y ecosistema tecnológico**  Cuando se establece la arquitectura, seleccionar el proveedor y el entorno digital y tecnológico es una tarea que se debe realizar teniendo en cuenta:   * Soporte técnico * Costos * Conocimiento de las herramientas.   Los entornos de productividad empresaria más populares y que aplican desde pequeñas empresas hasta grandes corporaciones son *Microsoft 360, Google Workspace, Amazon Web Service, IBM Services* |
| **Botón 8** | **Cultura y educación hacia la transformación digital**  Lo más importante para la transformación digital e integración de los sistemas de datos son las personas; de nada sirve invertir en alta tecnología o pagar a los mejores expertos técnicos, si los colaboradores, clientes y demás personas que hacen parte del ecosistema empresarial no tienen la educación y la cultura del cambio, induciendo a que las cosas se sigan haciendo de forma tradicional, aunque se tengan nuevas herramientas para registrar, optimizar y mejorar los procesos. La actitud de los colaboradores es un elemento fundamental para que la integración y cultura digital sea exitosa. |

# Calidad de los datos (*Quality data*)

| **Tipo de recurso** | Slider Imagen | |
| --- | --- | --- |
| **Introducción** | La calidad de los datos (DQ) se define como la facultad de los datos para el objetivo definido de un usuario u organización. Es un término subjetivo ya que el concepto de calidad podría ser relativo a los estándares definidos por las expectativas de las organizaciones o usuarios. (Gawande, 2020) | |
| El objetivo final de los datos es brindar conocimiento del negocio, de una manera técnica y fiel a la realidad, en una palabra, los reportes deben ser siempre **CONFIABLES** | | Imagen: 228131\_i109  **Imagen sugerida:** representación de confiabilidad.  <https://www.freepik.es/foto-gratis/apreton-manos-socios-comerciales-reunion-grupo-que-hace-inversion-proyecto-primer_3955349.htm#query=confiabilidad&position=0&from_view=search> |
| Una unidad diminuta, equivalente a elementos microscópicos en medio de los océanos de información es un único registro, en uno de los campos, de alguna tabla que componga el sistema (por ejemplo, FECHA DE NACIMIENTO) | | Imagen: 228131\_i110  **Imagen de muestra:** representación de conceptos campo – registro y dato.  <https://formacion.intef.es/pluginfile.php/37661/mod_imscp/content/1/tablas.html> |
| La colección de estos registros diversos debe tener desde el momento de capturar los datos una programación de validación que garantice que el registro de datos tenga aspectos como:   * Tipo de dato según la naturaleza del registro * Formato uniforme aceptado por todo el sistema * Dato válido según las reglas de negocio y naturaleza del proceso, entre otros. | | Imagen: 228131\_i111  **Validación de datos**  **Imagen de ejemplo**  **Tomada de:** [**https://www.freepng.es/png-flx02s/**](https://www.freepng.es/png-flx02s/) |
| La calidad de los datos toma aún mayor fuerza para aplicar validaciones con rigurosidad cuando se habla de ecosistemas de datos, donde convergen diversas infraestructuras y se interrelacionan datos compartidos. | | Imagen: 228131\_i112  **Datos compartidos en un ecosistema**  **Imagen de muestra**  Link: [https://www.freepik.es/vector-gratis/grafico-analisis-programacion-web-ilustracion-vector-estilo-plano\_10603911.htm#query= ecosistema%20de%20datos%20compartidos%20clientes%20proveedores%20sedes&position= 1&from\_view=search&track=ais](https://www.freepik.es/vector-gratis/grafico-analisis-programacion-web-ilustracion-vector-estilo-plano_10603911.htm#query=%20ecosistema%20de%20datos%20compartidos%20clientes%20proveedores%20sedes&position=%201&from_view=search&track=ais) |

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| **Inconvenientes derivados de mala calidad de los datos:**  Cuando los sistemas de información no son capaces de digitalizar eficaz y eficientemente la realidad del negocio, capturándola con el nivel adecuado de detalle necesario, almacenar dichos datos garantizando que no se produzcan pérdidas sintácticas ni semánticas, procesarlos de acuerdo con las reglas de negocio o mostrar los resultados de los análisis a los usuarios, entonces se producen “no-conformidades” en el ciclo de vida de los datos. (Velthuis, 2019) | |
| **Imagen sugerida que representa el concepto de no conformidades.**  <https://iso4docs.com/5-no-conformidades-mas-comunes-en-auditoria-iso-90012015/>  Imagen: 228131\_i113 | |

## Dimensiones de la calidad de los datos

| Cuadro de texto |
| --- |
| Se puede definir, que el conjunto de datos presenta calidad en los datos si cumplen las siguientes seis dimensiones;   * Exactitud * Completitud * Consistencia * Singularidad * Disponibilidad * Validez   Sin embargo, esta clasificación no está universalmente aceptada. En este componente formativo se adicionarán cuatro dimensiones *Data Quality* más:   * Actualizado * Conformidad * Integridad * Precisión |

**Ver Anexo2: CF10\_Dimensiones Quality**

| **Tipo de recurso** | Infografía interactiva Modal | |
| --- | --- | --- |
| **Texto introductorio** | Para una resolución integral del problema de calidad de datos, es necesario realizar un análisis de cada parámetro, lo que permitirá resolver todas las dudas que existen en el proceso y con ello reducir el riesgo de falla del proceso.  En la siguiente infografía podrá apreciar de manera más gráfica las diez dimensiones medibles que definen la calidad de los datos o Data Quality. | |
| **Figura 4**  *Dimensiones que definen la calidad de los datos*    Imagen de muestra, donde cada ítem tiene una imagen que representa el proceso.  Imagen: 228131\_i114 | | |
| **Código de la imagen** | 228131\_i114 | |
| **Precisión** | Grado en que los datos representan la realidad. | **Precisión**  **Imagen de muestra**  **Se puede construir como una línea de tiempo, pero con imágenes que representen cada proceso y al acercar el mouse, muestran la corta descripción:** |
| **Completitud** | Se establece como el porcentaje de datos poblados frente a la posibilidad de cumplimiento del 100 %. | **Completitud**  **Imagen de muestra** |
| **Consistencia** | Cercanía y uniformidad de los datos con otras tablas o un conjunto de datos de referencia. | **Consistencia**  **Imagen de muestra** |
| **Unicidad** | Registro único de cada evento. | **Unicidad**  **Imagen de muestra** |
| **Validez** | Proximidad del valor de los datos a valores predeterminados. | **Validez**  **Imagen de muestra** |
| **Puntualidad** | Exactitud entre el evento real y el registrado. | **Puntualidad**  **Imagen de muestra** |
| **Actualizado** | Captura de datos reales y actualizados. | **Actualizado**  **Imagen de muestra** |
| **Conformidad** | Los datos con los mismos atributos deben representarse en un formato y tipos de datos iguales. | **Conformidad**  **Imagen de muestra** |
| **Integridad** | Grado de coherencia con el que se implementan restricciones relacionales definidas entre dos tablas (tablas normalizadas). | **Integridad**  **Imagen de muestra** |
| **Precisión** | Se refiere al grado de las cifras redondeadas y segregadas. La precisión depende de los requisitos del sistema. | **Precisión**  **Imagen de muestra** |

## Datos maestros (*Master Data*)

| Tipo de recurso | Cajón de texto de color | |
| --- | --- | --- |
| Master Data es una arquitectura para administrar, centralizar, organizar, clasificar, localizar, sincronizar y enriquecer los datos según las reglas de negocio.  La gestión de datos maestros (MDM por sus siglas en inglés) se resume en un repositorio central que garantiza una única visión autorizada de la información y optimiza costos e ineficiencias causadas por los almacenamientos de datos dispersos. | | Imagen: 228131\_i115  Master Data  Imagen de muestra, tomada de: <https://blog.hubspot.com/website/master-data> |
| Imagen: 228131\_i116  Gestión de datos  Imagen de muestra  Link: <https://www.freepik.es/foto-gratis/concepto-gestion-datos-cloud-computing_17433093.htm#page=4&query=data&from_query=master%20data&position=14&from_view=search> | | La gestión de datos maestros apoya reportes de negocio mediante la ubicación exacta, vinculación y propiedades de entidades y de la información a través de productos, clientes, tiendas, ciudades, empleados, proveedores, activos digitales, y más. |
| MDM es un habilitador clave para proporcionar una vista única y confiable de la información empresarial crítica. Las fuentes de datos confiables ayudan a reducir los costos de integración de aplicaciones, mejoran la experiencia del cliente y generan información analítica accionable. (Stibo system MDM, 2019) Una solución de MDM procura superar algunos desafíos comunes en las organizaciones como:   | Mano con dedo índice apuntando a la derecha | Silos de datos dispersos y múltiples versiones de sus datos. | | --- | --- | | Mano con dedo índice apuntando a la derecha | Datos errados como resultado de ingresos manuales y datos no validados. | | Mano con dedo índice apuntando a la derecha | Costos en almacenamiento y seguridad en la información, tanto para su acceso, su conservación y disponibilidad | | | **Figura 5** *Características de MDM*    Características de MDM  Imagen propia  Imagen: 228131\_i117 |

# Arquitectura analítica de datos

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| Desde el enfoque técnico y el punto de vista estratégico existen diferencias marcadas entre el desarrollo de *software* y la analítica de datos.  Si bien se habla de tecnología y datos, el primero (el desarrollo de *software*) se enfoca a solucionar tareas de los procesos del negocio, es un sistema transaccional y la arquitectura tanto de procesos como de almacenamiento y diseño de las colecciones de datos enmarca una diferencia. Esta situación ha provocado en algunos casos un divorcio entre los programas de cómputo ya desarrollados y los sistemas de inteligencia de negocio y analítica; implicando gastos adicionales para las empresas en la integración, corrección y convalidación de datos entre varios sistemas de cómputo.  Desarrollo de Software Vs Analítica de Datos  Imagen propia elaborada en PPT  Imagen: 228131\_i118  Imagen original tomada de: <https://www.freepik.es/foto-gratis/dejame-pensar-grave-hombre-afeitar-barba-bigote-mantiene-dedo-indice-sien_13759451.htm> | |
| Antes de entrar a la explicación técnica es importante mencionar algunos conceptos teóricos, fundamentados en la estadística y las matemáticas para luego aplicar analítica de datos. | |

## Conceptos básicos estadísticos

| **Tipo de recurso** | Slider Presentación | |
| --- | --- | --- |
| **Introducción** | Generalmente en los materiales de estudio se insiste en que una buena decisión está basada en datos; efectivamente, en todas las áreas de las organizaciones se toman decisiones y por ello se requiere información de muchos tipos: finanzas, recursos humanos, mercadeo, clientes, logística, etc.  Para esto es necesario tener unos elementos básicos de estadística que se trabajarán a continuación; | |
| **Cuenta de datos financieros y de consumo:**  Para todo negocio es importante tener en cuenta datos financieros y de consumo.  Por ejemplo, en la siguiente imagen se aprecian los índices de precios, en los cuales se nota la inflación y consumo de la población en Colombia. Se nota que las bebidas alcohólicas y tabaco tuvieron estabilidad de precios a lo largo del mes de julio, siendo el calzado y prendas de vestir las que más tuvieron variación en sus precios, seguido de los alimentos.  Con esta información el negocio podría ir tomando decisiones y estableciendo enfoques comerciales; mirar cómo está su sector en el país y de esa manera ir orientando su análisis de negocio. | | *Índice de precios del consumidor*    **Índice de precios del consumidor (Banco de la república de Colombia, 2022)**  Imagen: 228131\_i119 |
| **Otro ejemplo de la utilidad de la estadística:**  Otro tema que podría interesar a los negocios, por ejemplo, es el uso de la tecnología y redes sociales de los colombianos. | | **Figura 7** *Estadísticas de uso de internet en Colombia 2021*  Estadísticas de Redes Sociales en Colombia 2021  **Imagen: Estadísticas uso de tecnología en Colombia 2021 (way2net, 2021)**  Imagen: 228131\_i120 |
| Si el negocio emplea la tecnología se podría verificar si en el país hay buena oportunidad en este aspecto y si se desea especificar más; por ejemplo, sobre los usuarios de redes sociales se podría revisar qué red social utilizan con mayor frecuencia.  Así, a partir de estos datos se decide qué redes sociales emplear más, adicional se evalúa el riesgo y oportunidad de invertir en negocios que usen la tecnología. | | **Figura 8** *Estadísticas de redes sociales en Colombia 2021*  Estadísticas de Redes Sociales en Colombia 2021  **Imagen: Estadísticas de Redes Sociales en Colombia 2021 (way2net, 2021)**  Imagen: 228131\_i121 |
| **Características de la Estadística**  Además, existen indicadores propios de las organizaciones, tales como la medición de desempeño de los colaboradores en áreas específicas, validación de datos financieros, análisis de los tiempos de procesos, comparación de sedes, definición de conocimiento específico y general del negocio.  Es por esto que la estadística toma un papel muy importante, para ello se mencionarán las principales características de la estadística:   * Objetivo y pregunta * Recopilación de los datos * Procesar los datos * Presentar los datos * Analizar los datos * Conclusiones y acciones | | **Figura 9** *Características de la estadística*    **Imagen: Características de la estadística**  **Construcción propia**  Imagen: 228131\_i122  [**https://image.shutterstock.com/image-photo/serious-business-man-trader-analyst-600w-1854622555.jpg**](https://image.shutterstock.com/image-photo/serious-business-man-trader-analyst-600w-1854622555.jpg) |

| **Tipo de recurso** | Acordeón tipo 1 |
| --- | --- |
| **Introducción** | Como principio fundamental la estadística trata con datos, donde se establecen: |
| **Figura 10** *Características de la estadística*  Tabla  Descripción generada automáticamente  Imagen: 228131\_i123 | |
| **Objetivos y preguntas**  Para empezar a hacer uso de la estadística es fundamental definir la intencionalidad de lo que se desea saber, comprobar o medir. En otras palabras; definir con claridad las preguntas que se desean responder. | |
| **Recopilación de datos**  Según la necesidad u objetivo del ejercicio estadístico, es necesario identificar si este existe o se debe diseñar la herramienta o instrumento para la recopilación de los datos. Verificar por ende si los sistemas de información tienen los datos requeridos e implementar encuestas o métodos de observación sistemática. | |
| **Procesar los datos**  Cuando se tengan los datos recopilados de manera sistémica, es importante tener claridad de qué hacer con ellos. Tener los datos no responde las preguntas requeridas si el proceso no está claro. Es fundamental interpretar y entender los datos para definir las fórmulas matemáticas y las mediciones para aplicar las operaciones a estos datos, siendo las operaciones estadísticas más comunes y simples: promedios, sumas, conteos, segmentación, porcentajes, variables, tiempo, etc. | |
| **Presentar los datos**  Presentar la información de manera amigable y fácil de entender que muestre con claridad las dimensiones de las medidas y las comparaciones, así como los puntos de referencia para poder determinar con claridad las cifras y su significado. La presentación de los datos se realiza a través de gráficos, tablas con varias dimensiones, medidas y segmentación. | |
| **Analizar los datos**  Este análisis que se presenta junto con la presentación de datos está basado en la interpretación y comprensión de la información que hay en los datos presentados; se definen también con medidas como desviaciones, comprobaciones estadísticas, proyecciones futuras, entre otras operaciones, esto para determinar el estado de lo que se ha trazado desde los objetivos y preguntas. | |
| **Conclusiones y acciones**  Si bien la estadística no se centra en las acciones, el fin es dar un conocimiento del negocio para que se use como insumo en la toma acertada de decisiones. | |

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| La estadística es entonces, un conjunto de métodos y teorías aplicadas a la recolección, descripción y análisis de datos, los cuales constituyen evidencia numérica para la toma de decisiones en condiciones de incertidumbre.  **Imagen Componentes de la estadística**  Imagen: 228131\_i124 | |

| **Tipo de recurso** | Acordeón tipo 2 |
| --- | --- |
| **Introducción** | **Tipos de estadística** |
| Imagen: 228131\_i125  **La estadística**  **Imagen de muestra**  Link: <https://www.freepik.es/foto-gratis/reporte-negocios-graficos-graficos-informes-negocios-monton-documentos-concepto-negocio_1275494.htm#query=estad%C3%ADstica&position=4&from_view=search> | |
| **Estadística descriptiva**  Hace referencia a los datos que se pueden representar a partir de tablas o gráficos.  En general describe fenómenos, por ejemplo, cuando se pregunta la edad a un grupo de personas; se podría realizar una tabla o un gráfico con estos resultados y definir medidas descriptivas como edad promedio, edades más frecuentes, etc. | |
| **Estadística inferencial**  Es la que se trabaja a partir de métodos a fin de sacar conclusiones, tomar decisiones, estimaciones o predicciones sobre una población o universo con base a datos de una muestra. | |

## Modelos matemáticos

| **Tipo de recurso** | Slider Presentación | |
| --- | --- | --- |
| **Introducción** | Para la aplicación de estadística inferencial, se tratarán a continuación algunos parámetros matemáticos que permiten la aplicación de este modelo estadístico. | |
| **Muestra Adecuada:**  Este tipo de estadística se aplica cuando la población o el universo objeto de estudio es muy grande e imposible de aplicar estadística descriptiva o aplicar instrumento a todo, o cuando no hay exactitud del número de población o universo.  Para estos casos se deberá aplicar la estadística inferencial donde se pueden realizar afirmaciones sobre una población basado en los resultados de una muestra.  Es importante no establecer divorcios entre la estadística descriptiva y la inferencial, pues la inferencial usa los datos descriptivos para llevar a conclusiones generales. | | **Figura 12** *Muestra adecuada*    Imagen: 228131\_i132  **Muestra adecuada**  **Imagen propia**  **Imagen: i228131\_62** |
| **Ejemplos de algunos usos de la estadística inferencial**   | **Gráfico de barras** | **Análisis de mercado:** existen empresas especializadas en *marketing* que realizan estudios que recogen datos relevantes a través de encuestas, *focus groups*, o de otros estudios previos y sacan deducciones tales como productos que prefiere la gente y los segmentos más sobresalientes. | | --- | --- | | **Gráfico de barras RTL** | **Epidemiología**: a partir de la toma de datos de ciertas zonas y patologías determinadas los especialistas en epidemiología y saneamiento público pueden llegar a conclusiones acerca de qué medidas de sanidad pública son necesarias para evitar que las enfermedades se expandan y cuidar más a la población. Las principales decisiones de gobiernos en la crisis COVID-19, fueron producto de la aplicación de tendencias tanto de estadística descriptiva como de estadística inferencial. |   Para iniciar un proceso de estadística inferencial, se debe partir de un supuesto u otro estudio como base para definir la hipótesis. | | Imagen: 228131\_i133  **Estadística inferencial**  **Imagen de muestra**  **Fuente:**  [**https://www.freepik.es/foto-gratis/concepto-reunion-trabajo-equipo-asunto-lanzamiento\_1235178.htm#query=an%C3%A1lisis%20de%20mercado&position=1&from\_view=search**](https://www.freepik.es/foto-gratis/concepto-reunion-trabajo-equipo-asunto-lanzamiento_1235178.htm#query=an%C3%A1lisis%20de%20mercado&position=1&from_view=search) |
| **Ejemplo de inferencia estadística**  Por ejemplo*,* para conocer la estatura media en Colombia sería una labor muy difícil realizar una encuesta a todos; por lo tanto, se emplea la estadística inferencial:   1. Lo primero es que de todos los colombianos (μ) se saca una muestra y una media de esa estatura (). 2. Para sacar la muestra no se puede ir simplemente un colegio y sacarles la medida a los jóvenes del equipo de voleibol, pues este segmento de personas será más alto que el promedio de la población. Además, la edad no sería tampoco representativa de todos. 3. Es importante que la muestra respete la diversidad de toda la población (todas las edades, ambos sexos casi por igual, estrato social, tipo de comunidad, etc.) | | Imagen: 228131\_i134  **Media de estatura**  Imagen de muestra tomada de: <https://www.freepik.es/vector-gratis/grafico-circular-personas_6550772.htm#query=estatura%20poblaci%C3%B3n&position=1&from_view=search&track=ais> |
| **Método para obtener la muestra representativa**  Para obtener la muestra de la población colombiana a fin de conocer la estatura media en Colombia, se procede a realizar un;  *Muestreo aleatorio simple:*  Cuando se aplica el instrumento para recoger datos a una muestra aleatoria se asume, por ejemplo, que se conoce de antemano que la media del tamaño de personas en Colombia es 170 cm con una desviación típica de 12.  μ=170 σ=12  **N (170,12)**  Lo anterior se puede representar en una curva como lo muestra la imagen; n la imagen, la curva representa que el máximo de personas tiene 170 cm. y de ahí, tanto los más altos (de 170 a la derecha) como los más bajos (de 170 hacia la izquierda) la población disminuye. Es decir, personas que midan 200 cm son pocas así como las que midan 140. | | **Figura 13** *Media de las medidas*    **Media de las medidas sabida de la población (referencia poblacional)**  Imagen propia  Imagen: 228131\_i135 |
| Ahora, se supondrá que se han tomado muestras de N personas.  Supóngase que N=225 personas  La media de esa muestra da la siguiente tabla:  **Tabla 2** *Datos de la muestra de N personas*    **Datos de la muestra de N personas**  **Tabla: 228131\_T2**  El gráfico resultante es el que se aprecia en la imagen. Como se puede notar es muy similar a la gráfica poblacional, por lo que las medias de las muestras escogidas forman una nueva distribución normal.  Población: μ=170  Muestra:  Calcular desviación:  N en la población  **N** en la muestra: | | **Figura 14** *Media de las medidas de la muestra*    **Media de las medidas de la muestra**  Imagen: 228131\_i136 |

## Gestión de datos

| **Tipo de recurso** | Pestañas o *tabs* horizontales | |
| --- | --- | --- |
| **Introducción** | Las estadísticas desde el entorno digital y tecnológico han apalancado los términos como analítica (*analytics*) y *Big* *Data*. En estas áreas se toman datos, se obtienen resultados y se visualizan los datos  Para gestionar los datos lo primero es saber desde qué enfoque se desea realizar esta gestión. A continuación, se describe el proceso: | |
| **Obtención de datos** | **Obtención de datos de los procesos**  • Las organizaciones actuales, tienen enfoques hacia los procesos, y todos los procesos deberían de ir asociados de manera directa o indirecta a los clientes; por lo que se dice que las organizaciones no dependen de personas sino de procesos y adicionalmente, tienen un enfoque hacia el cliente.  • En este sentido cada proceso o área de la organización, debe tener claridad no solo de sus funciones, sino, además de los objetivos y las metas, según las funciones en toda la cadena general de la empresa.  • Si bien algunas pymes con un enfoque tradicional consideran que no tienen necesidad de capturar datos de sus procesos, pues cada área tiene un responsable y basta con que la gerencia les consulte para saber cómo va el negocio; esto no es lo más adecuado; y menos cuando existen las posibilidades de utilizar la analítica y la inteligencia de negocios al alcance de todos.  • Otro caso particular entre muchas empresas es que usan los programas informáticos en algunas áreas, pero no se encuentran integrados o relacionados. Si bien usan tecnología digital para mejorar y optimizar tareas y procesos, ante los ojos de la gerencia y desde la mirada holística corporativa podría existir carencia de sinergia; aquí también en caso de que la gerencia requiera información sobre su negocio, deberá llamar o solicitar informes a cada líder para saber cómo marcha la organización.  • Un elemento indispensable para la gestión de los datos en las organizaciones es determinar área por área los métodos y tecnologías que usan para la captura de los registros o de los procesos o sucesos de cada departamento. | Imagen: 228131\_i137  **Obtención de datos**  Crear una imagen que simule la obtención de datos por diferentes fuentes.  Imagen de muestra. |
| **Requisitos captura de datos** | Algunos requisitos de captura de datos de todas las áreas deberán cumplir con los siguientes parámetros:  **Nota: si es posible insertar aquí un recurso de Pestañas o Tabs verticales para cada ítem, si no, representar la información en tabla:**   | **Descargar desde la nube** | 1. ***[Qué] Registrar Hechos:*** Todas las áreas deben registrar las tareas o acciones (compras, terminación de tareas, visitas, publicación de anuncios, etc.). | | --- | --- | | **Descargar desde la nube** | 1. ***[Cuándo] Registrar momentos:*** Una de las características de la analítica es llevar una historia, no es posible que exista una historia si no hay fechas. Es importante determinar las fechas de cada hecho. | | **Descargar desde la nube** | 1. ***[Quién] Registrar responsable:*** Cada tarea deberá tener un responsable, más cuando hay tareas entre diferentes áreas, de esta manera es posible identificar quién o en dónde se encuentran los embudos en los diferentes procesos del negocio. | | **Descargar desde la nube** | 1. ***[Cuánto] Registrar cantidades:*** La mayoría de los hechos, contienen un número que acompaña las características, ejemplo, una venta tiene detalles como cantidad de elementos y un monto de dinero, estas cifras siempre deberán registrarse; en otras tareas, como por ejemplo aprobar una solicitud, donde no hay un monto numérico sí contiene un estado, por lo que también debe registrar los estados de estas tareas. | | **Descargar desde la nube** | 1. ***[A quién o de quién] Registrar clientes o proveedores:*** Cada hecho o acción debe tener un cliente, si bien puede ser externo como las ventas también podría ser interno, por ejemplo, la aprobación de presupuesto, el cliente interno sería la dirección responsable de lo financiero; todos los hechos deben tener a quién o para quién se le hace la acción. |   No basta con tener registrados estos cinco mínimos componentes de las tareas más relevantes para que los datos muestren una radiografía clara sobre cada proceso. Además, es importante tener claridad sobre la calidad de los datos para que se puedan integrar y tener plena confianza. | **Figura 15** *Parámetros de captura de datos*    Imagen: 228131\_i138  **Parámetros de captura de datos**  **Imagen propia elaborada en PPT**  **Imagen original:** [**https://www.freepik.es/foto-gratis/exito-traves-trabajo-equipo-formacion-equipos\_3687011.htm#query=gente%20y%20elementos&position=4&from\_view=search**](https://www.freepik.es/foto-gratis/exito-traves-trabajo-equipo-formacion-equipos_3687011.htm#query=gente%20y%20elementos&position=4&from_view=search) |
| **Integración de datos** | **Integración de datos**  Ocasionalmente no existe un sistema que sirva para todo cuando se tienen identificadas las fuentes de datos de cada área se procede a extraer estos datos.  Existen diversidad de fuentes de datos; se deben diseñar metodologías para que los datos tengan una copia óptima en sistemas de almacenamiento centralizados que garanticen la calidad de estos y estén disponibles para la realización de reportes de todas las áreas. | Imagen: 228131\_i139  Integración de datos  **Crear una imagen que represente la integración de datos.**  **Imagen de muestra, tomada de:** [**https://cio.com.mx/tres-elementos-a-menudo-se-omiten-en-la-integracion-datos/**](https://cio.com.mx/tres-elementos-a-menudo-se-omiten-en-la-integracion-datos/) |
| **Diseño de reportes** | **Diseño de reportes**  El objetivo es que todos los datos recolectados, copiados y optimizados sean consumidos por el nivel de decisión; no necesariamente son los gerentes, los empleados también deben acceder a sus gestiones y resultados en gráficos y *dashboards* que den cuenta de su propio rendimiento, esto permite una autoevaluación y la toma de decisiones sobre sus propias funciones en la empresa.  En términos generales, la gerencia ya no tendrá que llamar a los líderes de cada proceso para que le den explicación sobre diferentes situaciones ya que tendrá acceso a las cifras en sus teléfonos celulares, en cualquier lugar y a cualquier hora. Lo que permite tener empresas más eficientes, con capacidad de reaccionar más pronto y tomar mejores decisiones. | Imagen: 228131\_i140  Diseño de reportes  **Crear una imagen que represente la obtención de reportes en dispositivos móviles.**  **Imagen de muestra, tomada de:** <https://www.freepik.es/foto-gratis/concepto-rendimiento-analisis-informacion-datos_16469936.htm#query=reportes%20en%20dispositivos%20m%C3%B3viles&position=6&from_view=search> |

**Anexo 4 – Infografía 2**

# Generalidades: Herramienta inteligencia de negocio

| Tipo de recurso | Cajón de texto de color |
| --- | --- |
| En términos generales, la inteligencia de negocio no es una tecnología o una serie de herramientas ya establecidas, es decir, todo depende de qué tan grande y compleja sea la organización y qué cantidad de datos fluyen en los sistemas.  Por lo que la inteligencia de negocio podría gestionarse desde la aplicación de Excel u Hojas de cálculo de Google hasta usar herramientas especializadas con proveedores, entre otros, como *Amazon Web Services* o *Microsoft* con sus opciones de *Azure* que incluso podrían ser tan sencillas como complejas y costosas según la organización. | |
| Existen variadas herramientas para la inteligencia de negocios, en este componente formativo se trabajarán en una de las más populares de la marca Microsoft denominada *Power* BI. | |

## Instalación

| Cuadro de texto |
| --- |
| Microsoft distribuye de manera gratuita la aplicación para conectar datos y generar reportes, esta aplicación se encuentra bajo el ecosistema de productividad Microsoft 365, la cual, se cobra solamente para algunas características extras, pero para efectos del aprendizaje, la versión descargable es suficiente y muy completa.  Para su descarga se debe abrir en el navegador el centro de descarga de Microsoft,  <https://www.microsoft.com/es-ES/download> luego dar clic en la opción *Office*, y luego en *Microsoft Power BI Desktop*  Imagen: 228131\_i141  *Imagen: Logo de Power BI*  *Imagen de muestra, tomada de: https://teorema-rd.com/producto/curso-da-100t00-a-analyzing-data-with-powerbi* |

| Cuadro de texto |
| --- |
| En el siguiente video podrá aprender cómo se realiza la instalación del componente. |

**Ver Anexo 5 – CF10\_ guion Video temático**

## Reportes (*Data Source*)

| **Tipo de recurso** | Rutas /Pasos Horizontal | |
| --- | --- | --- |
| **Introducción** | Este tipo de aplicaciones pueden ser muy fáciles y básicas, aunque también se podrán desarrollar *dashboard* tan complejos y robustos según el nivel de conocimiento de las herramientas.  En este apartado, se tratarán los reportes de manera resumida: | |
| **Paso 1** | **Aplicaciones**  Las aplicaciones de visualización de datos profesionales; por lo general no hacen todo en una misma aplicación, sino que trabajan varios programas que se pueden conectar entre sí. | Imagen: 228131\_i142  **Visualización de datos**  **Imagen de muestra que representa diversas formas de representación de datos.**  Tomada de: [https://www.freepik.es/vector-gratis/sistema-i nterfaces-manipulacion-directa-maquina-humana-holografica-composicion-isometrica-comunicacion-analisis-transferencia-datos-grandes\_7380081.htm#query=visualizaci%C3%B3n%20de%20datos&position=4&from\_view=search](https://www.freepik.es/vector-gratis/sistema-i%20nterfaces-manipulacion-directa-maquina-humana-holografica-composicion-isometrica-comunicacion-analisis-transferencia-datos-grandes_7380081.htm#query=visualizaci%C3%B3n%20de%20datos&position=4&from_view=search) |
| **Paso 2** | **Power BI**  Las fuentes de datos de *Power BI*, son diversas; de manera nativa, el programa tiene decenas de opciones que son las más comunes, incluso es tan compatible que, si existiera una fuente a partir de un desarrollo no comercial o poco común, daría la posibilidad de programarlas y crearlas.  Existen diversas fuentes, desde archivos, pasando por múltiples motores de bases de datos, plataformas, marcas, servicios en línea, etc. | **Figura 16** *Fuente de datos Power BI*    **Imagen:** Fuentes de datos Power BI  Imagen propia  Imagen: 228131\_i143 |
| **Paso 3** | **Microsoft Excel**  Entre las opciones de fuentes de datos, está MS-Excel.  Cada archivo es un libro y cada hoja del libro se toma como una tabla, de tal modo que si hay conjunto de datos dentro de Excel, el programa lo asume como tablas independientes. | **Figura 17** *Hoja de Excel como tabla*    **Imagen:** Hoja de Excel como tabla  Imagen propia  Imagen: 228131\_i144 |
| **Paso 4** | **Implementación**  Cuando se elige la tabla, se da clic en cargar y ya los datos estarán disponibles para empezar a desarrollar los reportes.  En el área de trabajo se visualizarán las columnas o campos del archivo fuente. | **Figura 18** *Área de trabajo*    **Imagen:** Área de trabajo  Imagen propia  Imagen: 228131\_i145 |

## Desarrollo y representación gráfica de datos

| **Tipo de recurso** | Slider Presentación | |
| --- | --- | --- |
| **Introducción** | Luego de tener la aplicación instalada y con la fuente de datos clara lo que sigue es crear informes: | |
| Suponiendo que ya se han cargado datos la tabla cargada contiene campos, por ejemplo, ciudades; se arrastra entones el campo al área de trabajo. | | **Figura 19** *Campos de la tabla*    **Imagen:** Campos de la tabla  Imagen propia  Imagen: 228131\_i146 |
| Se visualizan entonces los datos del campo arrastrado. | | **Figura 20** *Datos del campo ciudad*    **Imagen:** Datos del campo ciudad  Imagen propia  Imagen: 228131\_i147 |
| Ahora bien si la tabla tiene la producción por año en cada ciudad es interesante visualizar dichos datos, para diferentes años, por ejemplo, en este caso, se puede acceder a los datos de producción de los años 2018 y 2019. | | **Figura 21** *Producción de los años 2018-2019*    **Imagen:** Producción de los años 2018-2019  Imagen propia  Imagen: 228131\_i148 |
| Los datos para los años 2018 y 2019 en cada ciudad, se observan así: | | **Figura 22** *Producción en cada ciudad para los años 2018-2019*    **Imagen:** Producción en cada ciudad para los años 2018-2019  Imagen propia  Imagen: 228131\_i149 |
| Ahora, si bien para algunos datos las tablas son difíciles de sustituir, sería más amigable usar otros recursos gráficos. *Power* BI cuenta con muchos tipos de visualizaciones. | | **Figura 23** *Tipos de visualización en Power BI*    Tipos de visualizaciones de Power BI  Imagen propia  Imagen: 228131\_i150 |
| Para el ejemplo de las ciudades, se seleccionó “gráfico de barras agrupadas”  El resultado es la visualización más óptima para revisar la productividad por ciudades de cada año, según se observa en la imagen: | | **Figura 24** *Gráfico de barras agrupadas*    **Gráfico de barras agrupadas**  **Imagen propia**  Imagen: 228131\_i151 |
| También usando otros campos de datos se pueden combinar con otras visualizaciones, en este caso al tratarse de ciudades, se puede implementar la visualización de mapas. | | **Figura 25** *Gráfico de barras agrupadas con mapas*    **Gráfico de barras agrupadas con mapas**  **Imagen propia**  Imagen: 228131\_i152 |
| Los reportes tienen muchas posibilidades adicionales, con conjuntos de datos y otras tablas relacionadas, con filtros y configuración de condiciones, colores, formatos, etc., lo que hace de *Power BI* una herramienta muy poderosa. | | Imagen: 228131\_i153  **Reportes de datos**  **Imagen sugerida**  Link: https://img.freepik.com/free-vector/data-report-illustration-concept\_114360-883.jpg?w=2000 |

**Síntesis**

| **Tipo de recurso** | Síntesis |
| --- | --- |
| Gestión de analítica de datos  Síntesis: análisis y conocimiento de datos enfocada a la analítica | |
| **Introducción** | En el siguiente mapa conceptual se podrán visualizar los principales conceptos trabajados con respecto al análisis de datos para analítica: |
| **Figura 26** *Análisis de datos para analítica*    Nombre: Cmap Análisis de datos para analítica  Imagen: 228131\_i154  Fuente online: <https://cmapscloud.ihmc.us/viewer/cmap/1YCNDSHLR-MMWXK8-FFG320> | |

**Material complementario**

| Tipo de recurso | Material complementario | | |
| --- | --- | --- | --- |
| **Tema** | **Referencia APA del material** | **Tipo** | **Enlace** |
| Aspectos a tener en cuenta sobre los ecosistemas de datos. | IBM Services (2022). *Servicios y consultoría*. IBM Services. | Servicios empresariales | <https://www.ibm.com/co-es/services> |
| Aspectos a tener en cuenta sobre los ecosistemas de datos. | Microsoft 365 (2021). *Portafolio de servicios de Microsoft 365*. Microsoft. | Portafolio de servicios | <https://goo.su/rfbWJ> |
| Aspectos a tener en cuenta sobre los ecosistemas de datos. | AWS Educación (2020) *Informática en la nube con AWS*. AWS. | Manual Computación en la nube de AWS | <https://aws.amazon.com/es/what-is-aws/?nc2=h_ql_le_int> |
| Aspectos a tener en cuenta sobre los ecosistemas de datos. | Google (2022) *Google Workspace. Características* Google *Workspace*. Google Workspace. | Portafolio de servicios | <https://workspace.google.com/intl/es-419/features/> |
| Inconvenientes derivados de mala calidad de los datos. | Greenfield, D. (2022) *Cinco causas de la mala calidad de los datos*. Boletín Mundo PMMI. | Documento web | <https://www.mundopmmi.com/automatizacion/adquisicion-de-datos/article/22018270/cinco-causas-de-la-mala-calidad-de-los-datos> |
| Conceptos básicos estadísticos. | SENA. [Ecosistema de Recursos Educativos Digitales SENA] (2022). *Herramientas de información estadística básica* [Video]. Youtube. | Video | <https://youtu.be/AW1LM-d0YWE> |
| Generalidades: Herramientas de inteligencia de negocio | Makesoft Technologies (2017) *Power BI Desktop: Tutoriales y Manuales en español*. Makesoft Technologies. | Tutorial Power BI | <https://www.makesoft.es/power-bi-desktop-tutoriales-y-manuales-en-espanol/> |

**Glosario**

| **Tipo de recurso** | Glosario |
| --- | --- |
| Atributos de entidades | las entidades se refieren a los objetos o hechos de datos en un sistema (ejemplo, clientes, ventas, etc.) y los atributos a sus características (ejemplo, edad, nombre, fecha) |
| *Bits* | es la unidad mínima de información, desde el punto de vista digital, es lo equivalente al átomo en la biología. Solo puede albergar dos datos posibles: 0 o 1 |
| *Bytes* | es el equivalente a una palabra, en términos digitales un byte equivale a 8 bits. |
| Convalidación de datos | ocasionalmente en los procesos ETL se deben unificar datos que, aunque estén de manera distinta significan lo mismo, por ejemplo, el campo sexo, para algunos sistemas se denomina género, por lo tanto, se debe convalidar y dejar un solo nombre, igual ocurre con los registros, para masculino es igual hombre, M, masculino, etc. |
| *Dashboards* | tableros que representan datos e informes. Representan en una sola vista simple una cantidad importante de datos; la idea es visualizar de manera fácil datos complejos que vienen desde los sistemas de información. |
| *Focus groups* | método para la investigación o estadística cualitativa, que consiste en reunir un grupo de personas para una entrevista, donde se expresan percepciones y opiniones de productos o servicios. Esto sirve como estudio para las marcas y productos. |
| Granularidad | se refiere a la resolución o detalle de los datos, el mejor ejemplo de granularidad es la referencia a los datos; se puede granular por año, por semestre, trimestre, mes, semana, día u horas. |
| Reglas de negocio | en el desarrollo de sistemas de información se refiere a las condiciones particulares y el comportamiento que deben asumir los algoritmos programados, para que los resultados sean los esperados por el proceso o negocio. |
| Silos de datos | es el almacenamiento de grandes cantidades de datos que por lo general son diversos; se usan para proteger los datos, pueden contener datos procesados o simplemente datos registrados como respaldo. |
| Validación | es la revisión de un dato para que cumpla con reglas o restricciones, así como para garantizar que los formatos de los datos sean los correctos y mejorar la calidad de los mismos. |

**Referencias bibliográficas:**

| **Tipo de recurso** | Bibliografía |
| --- | --- |
| Banco de la República de Colombia. (julio de 2022). *Inflación total y meta*. Estadísticas Banco de la República. <https://www.banrep.gov.co/es/estadisticas/inflacion-total-y-meta> | |
| Gawande, S. (2020). *6 Dimensions of Data Quality, Examples, and Measurement*. *iCEDQ Torana INC.* <https://icedq.com/6-data-quality-dimensions> | |
| Greenfield, D. (25 de enero de 2022). *Cinco causas de la mala calidad de los datos.* Mundo PMMI. <https://www.mundopmmi.com/automatizacion/adquisicion-de-datos/article/22018270/cinco-causas-de-la-mala-calidad-de-los-datos> | |
| Microsoft. (2022). *Tipos de datos y funciones.* Documentación de SQL. <https://docs.microsoft.com/es-es/sql/t-sql/functions/date-and-time-data-types-and-functions-transact-sql?view=sql-server-ver16> | |
| RSM. (11 de agosto de 2020). *Generación de valor desde los ecosistemas digitales.* RSM Global. <https://www.rsm.global/colombia/es/ideas/consulting-insights/generacion-de-valor-desde-los-ecosistemas-digitales> | |
| Stibo Systems. (octubre de 2019). *¿Qué es la gestión de datos?* Stibo Systems Master Data Management. <https://www.stibosystems.com/es/what-is-master-data-management> | |
| Velthuis, M. P. (2019). *Calidad de datos.* Ediciones de la U. <https://www-ebooks7-24-com.bdigital.sena.edu.co/?il=9094> | |
| Way2net. (13 de noviembre de 2021). *Estadísticas de Redes Sociales en Colombia 2021*. Agencia de Marketing Digital Way2net. <https://www.way2net.com/2021/11/estadisticas-de-redes-sociales-en-colombia-2021/> | |