**Datos de identificación del programa de formación**

| PROGRAMA DE FORMACIÓN | Asistencia para la inteligencia empresarial |
| --- | --- |

| COMPETENCIA | 220501114. Sistematización de datos masivos de acuerdo a técnicas de analítica de datos y herramientas tecnológicas. | RESULTADOS DE APRENDIZAJE | 220501114-01. Organizar la información a sistematizar de acuerdo con técnicas de análisis.  220501114-02. Realizar el afinamiento de los datos de acuerdo con las herramientas informáticas seleccionadas. |
| --- | --- | --- | --- |

| NÚMERO DEL COMPONENTE FORMATIVO | 6 |
| --- | --- |
| NOMBRE DEL COMPONENTE FORMATIVO | Procedimiento de analítica de datos |
| BREVE DESCRIPCIÓN | El fundamento clave del componente es la gestión de datos empresariales. El proceso analítico de los datos es el elemento de entrada: tener clara la procedencia, control y protección. Se enfatiza en el alistamiento de los datos, conociendo la conformación de las bases de datos, así como también los procesos de extracción, transformación y preparación óptima para el análisis empresarial. |
| PALABRAS CLAVE | Datos, análisis, procedimientos, bases de datos, gestión. |

| ÁREA OCUPACIONAL | 1 - FINANZAS Y ADMINISTRACIÓN |
| --- | --- |
| IDIOMA | Español |

# **TABLA DE CONTENIDOS**

**1. La analítica y resguardo de los datos en la empresa**

**2. Transformación de datos para uso empresarial**

2.1 Herramientas de gestión de datos

2.2 Análisis exploratorio de datos y alistamiento de los sets de datos

**3. Las bases de datos**

3.1 Tipos de bases de datos

3.2 Procesos ETL para el manejo de bases de datos

**INTRODUCCIÓN**

| Cuadro de texto |
| --- |
| Apreciado aprendiz, bienvenido a este componente formativo, donde se muestra la secuencia desde el momento en que se requiere un dato en la empresa, hasta la etapa de difusión y uso de los datos en los procesos, cumpliendo con su almacenamiento para tratamientos futuros. Se realizará la descripción de las bases de datos, desde sus distintos orígenes y clasificaciones, además de su apropiación, dependiendo de las necesidades de la empresa y haciendo énfasis en los procesos estandarizados de extracción, transformación, preparación y distribución de datos.  En el siguiente video conocerá, de forma general, la temática que se estudiará a lo largo del componente formativo. |

**GUION DE VIDEO INTRODUCTORIO**

| **Tipo de recurso** | Video motion | | | |
| --- | --- | --- | --- | --- |
| **NOTA** |  | | | |
| **Título** | Introducción | | | |
| **Escena** | **Imagen** | **Sonido** | **Narración** | **Texto** |
| **1** | Composición isométrica de servicios en la nube con imágenes conceptuales de cápsulas de almacenamiento de elementos de red y pequeños personajes humanos ilustración vectorial |  | Bienvenido a este componente formativo, donde se estudiará la temática relacionada con el proceso para examinar conjuntos de **datos**.  Se aprenderá sobre análisis de datos y cómo, en las entradas, los datos provienen de las necesidades de producción enfatizando en: la demanda, la disponibilidad de recursos de los indicadores meta, la cantidad de maquinaria disponible, de los proveedores, quienes suministran insumos y materias primas e incluso datos, de la competencia, que permite orientarse a producir con mayor calidad y lograr alta competitividad. | Analítica de datos  Entradas  Necesidades de producción |
| **2** | Los centros de datos son centros de datos para servir aplicaciones e información comercial de apoyo. tecnología de computación en la nube para análisis de negocios, analítica, investigación, estadística estratégica, planificación, marketing. |  | Las empresas funcionan con datos; ameritan datos para su funcionamiento desde las entradas, y generan nuevos datos en la salida de cada proceso. Es así como los datos cualitativos o cuantitativos, dinamizan a la organización.  Se aprenderá, cómo los nuevos datos indican si lo que se propuso, se logró. Se ahonda en la manera en que esos datos pasan a un proceso de analítica de datos, el cual conlleva a procesos de retroalimentación y toma de decisiones para el mejoramiento progresivo y permanente de los procesos. | Datos de entrada y de salida  Proceso de analítica de datos |
| **3** | Cv de empleados, cv de candidatos. trabajadores corporativos, id de estudiantes aislar elemento de diseño plano. solicitudes de empleo, avatares, ilustración del concepto de información personal |  | Puede convertirse en experto en lo referente a la generación de datos. Será protagonista de la toma de decisiones en su empresa; pero, nos preguntamos: ¿cómo intervienen los datos en la toma de decisiones?  Los datos que ingresan o se generan del sistema empresarial, orientan en la toma de decisiones para: cambios de estrategia, mejoras de procesos y aprovechamiento de oportunidades.  Aprenderá a ser líder de la gestión oportuna de datos, apoyada en adecuados sistemas de información, los cuales no siempre deben ser altamente sofisticados, sino más bien efectivos. | ¿Cómo intervienen los datos en la toma de decisiones?  Cambios de estrategia  Mejoras de procesos  Aprovechamiento de oportunidades |
| **4** | Business brainstorming graph chart informe de datos concepto |  | Es necesario recalcar cómo gestionar los datos internos, lo que implica reconocer cuáles son las variables claves de los procesos que afectan la productividad, la calidad y la competitividad; este sería el caso de cantidad de: unidades producidas, clientes satisfechos, productos que se poseen por sobre los de la competencia, volúmenes de venta, magnitud de desperdicios, unidades producidas con inconformidades, costos de mantenimiento, entre otros, dependiendo del tipo de empresa y del tipo de proceso. | Cómo gestionar los datos internos  Productividad  Calidad  Competitividad |
| **5** | Personas dibujadas a mano recopilando datos con estadísticas. |  | Este componente le aportará conocimiento para fortalecer sus habilidades en registro, almacenamiento y procesamiento de datos; a convertirlos en información que pueda ser utilizada por la gerencia y que, a su vez, provea datos para otros procesos de la empresa. En general, comprender de dónde provienen los datos utilizados por su organización.  Normalmente, los entes externos gubernamentales, ofrecen datos como la inflación, precios estándar, demanda potencial, competencia de productos, niveles de crecimiento o disminución del mercado, todos fundamentales para la planificación de la producción o la prestación de los servicios. | Registro  Almacenamiento  Procesamiento |
| **6** | Cumplimiento norma ley y regulación interfaz gráfica para política de calidad empresarialFondo de tecnología de innovación para finanzas empresariales |  | Existen datos en normas, decretos o leyes, o en organizaciones empresariales, que viabilizan los procesos estratégicos, administrativos y operativos.  Desde ya, tenga presente: una vez clarificados los datos necesarios para el funcionamiento del sistema empresarial, se hace necesario, en la empresa, el uso de la tecnología, como mínimo la más básica, para la extracción, ordenamiento, análisis y generación de información.  ¡Le deseamos éxito en este proceso de aprendizaje! | Uso de tecnología  Extracción  Ordenamiento  Análisis  Generación de información |
| **Nombre del archivo** | 134104\_v1 | | |  |

**DESARROLLO DE CONTENIDO**

**1. La analítica y resguardo de los datos en la empresa**

| Cuadro de texto |
| --- |
| Para comenzar, debemos comprender el análisis de los datos como un proceso sistemático que amerita de los datos como insumo, pero además, de procesos de pensamiento y de herramientas tecnológicas para poder extraer de los datos la información precisa, certera y oportuna que amerita la empresa para el mejoramiento de sus procesos o, en general, para cualquier tipo de toma de decisiones, sea general o específica. Los datos que se obtienen de los procesos de la empresa, o de fuentes externas, luego del proceso de análisis, pasan a una etapa de aseguramiento, lo que se constituye en el resguardo de los datos bajo los principios y necesidades de la empresa. Colaboración de equipo de negocios discutiendo el análisis de trabajo con datos financieros y el gráfico de informes de crecimiento de marketing en equipo, presentación e intercambio de ideas para la planificación de estrategias con beneficio de la empresa.  Ahora, la analítica de datos es un proceso sistemático que requiere de diversos elementos como insumo (los datos), de los procesos de observación de la persona quien realiza los análisis, y de las herramientas tecnológicas en las cuales se apoya, para tal fin. |

| Cuadro de texto |
| --- |
| En el marco de la analítica de datos, el análisis corresponde al proceso en el cual, luego de obtenidos los datos, se estructuran, agrupan, comparan, diferencian, se miden, se relacionan y finalmente sintetizan, con respecto a los intereses de una variable o realidad propia de la empresa, en la cual hacen falta esos datos para que un proceso gerencial, administrativo u operativo funcione, sea una máquina que amerite datos de entrada, o el establecimiento de indicadores de control de procesos o de productividad, o decisiones de operaciones de la empresa que se realizan a partir de los datos. Aunque el análisis es un proceso de la mente humana, las tendencias tecnológicas han desarrollado *software* que cumplen funciones integrales de apoyo al trabajo de las personas. |

| **Tipo de recurso** | | Pestañas o tabs Verticales |
| --- | --- | --- |
| **Introducción** | | En la analítica de datos encontramos: |
| Análisis de datos, visualización de bases de datos. Manos humanas en el teclado de la laptop. Persona que trabaja ordenando información usando mapa digital de la mente. Datos de salida de entrada, gráficos, análisis, ilustración vectorial de infografía | | |
| **Analítica de entrada** | La analítica de datos tiene una causa, la cual es la obtención de información sistematizada, debido a que en la empresa se requiere información para poner en marcha los procesos de producción o de servicio, para mejorarlos, para crear nuevos procesos o, sencillamente, para conocer el comportamiento de la oferta y la demanda de un producto, o las tendencias de los factores económicos de una región que pueden afectar a la empresa, entre otros. Lo importante es saber que el funcionamiento de los procesos, amerita la entrada de datos. | |
| **Analítica de salida** | Si se observan los datos como consecuencia, entonces pasan a tomar un papel clave como salida o resultados. De la operación de los procesos de manufactura, o de la prestación de los servicios, se obtienen datos de salida, sean volúmenes de venta, número de cartera de clientes atendidos, cantidades producidas, unidades fuera de especificaciones, niveles de rotación de recurso humano, cantidad de nuevos productos o servicios, crecimiento de la competencia, entre otros, que pasan luego al proceso de retroalimentación para el análisis de los resultados globales del sistema empresarial y el establecimiento de correctivos. | |

| Cuadro de texto |
| --- |
| Los datos de la empresa pueden provenir de diferentes fuentes. |

| **Tipo de recurso** | Pestañas o tabs horizontales | |
| --- | --- | --- |
| **Introducción** | Las cuales pueden ser: | |
| **Fuentes  internas de  datos** | Son conocidas como fuente primaria de datos y son específicas de los procesos, producidos por la empresa y por particulares para cada proceso. Luego que se ejecuta un proceso y se obtienen resultados, esos datos empiezan a funcionar como “datos históricos”, los cuales se van acumulando en bases de datos, y sirven para hacer comparaciones de mejora o de bajas de desempeño de la propia organización. Hay datos internos que se consideran “estándares”, y corresponden a las metas iniciales que, en un período de actividad empresarial, se constituyen en resultados esperados; por ejemplo: metas de venta generales en cantidades monetarias, metas de aumento de cartera de clientes, metas de producción, metas de disminución de índices de fallas en las maquinarias, entre otros. | Icono de concepto de fusión. Ilustración de línea delgada de idea abstracta de expansión externa. Fusión de dos empresas en una. Alianza empresarial. Dibujo de color de contorno aislado del vector. Trazo editable |
| **Fuentes externas de datos** | Los datos externos provienen de fuentes secundarias, es decir, ajenas al sistema empresarial, pero son datos que la empresa amerita para desarrollar sus procesos. Se puede tratar de datos esenciales o de datos de apoyo. Los datos esenciales son, por ejemplo, los costos de los insumos y materias primas en el mercado, la cantidad y ubicación de proveedores, la cantidad y ubicación de la competencia, los precios de la competencia; todos estos datos se ameritan para las decisiones de producción y mercadeo. Otros datos externos que sirven de apoyo para los cálculos o decisiones de procesos internos son, por ejemplo, las estadísticas nacionales de economía como inflación, índices de precios al consumidor, datos en cuanto a normativas, leyes y reglamentos de direccionamiento de producción, cambios en las preferencias de consumo de productos y servicios que realizan empresas especializadas, entre otros. | Icono de concepto turquesa interno y externo. Ilustración de línea delgada de idea abstracta de administración de comunicaciones de proyectos. Dibujo de contorno aislado. Trazo editable. Fuentes Arial, Myriad Pro-Bold utilizadas |

| Cuadro de texto |
| --- |
| Asimismo, los datos de la empresa pueden ser cuantitativos o cualitativos. Conozcamos cada uno de ellos: |

| **Tipo de recurso** | Acordeón tipo 1 |
| --- | --- |
| Concepto de tableros estratégicos, diseño de iconos vectores, símbolo de estrategia de negocio, firma de plan de marketing, ilustración de gestión de administración y de material operativo | |
| **Datos cualitativos**  Corresponden a elementos que no pueden ser medidos o sometidos a estadística por su naturaleza de “cualidad” o caracterización de algún aspecto o fenómeno de análisis. Se basan casi siempre en opiniones de las personas, pero son fundamentales en las empresas para los procesos de análisis. Un dato cualitativo puede ser el producto de una entrevista realizada a un empresario para una nueva toma de decisión en torno a la empresa; también puede tratarse de un artículo del **Código de Comercio** que restringe o viabiliza un proceso de producción o una actuación empresarial, o la opinión de los clientes con respecto al servicio que están recibiendo y los puntos de mejora que consideran necesarios para la satisfacción. La empresa puede generar datos cualitativos como, por ejemplo, la opinión del personal en cuanto al clima laboral o los factores de motivación y, a su vez la empresa, puede utilizar datos cualitativos de fuentes externas. | |
| **Datos cuantitativos**  Son aquellos datos numéricos que pueden procesarse con herramientas tecnológicas de apoyo a través de la estadística y conllevan al establecimiento de tendencias en el comportamiento de los procesos empresariales. La mayoría de los datos en las empresas son cuantitativos y la norma básica de la mejora de la productividad empresarial dice que lo que se mide se puede conocer, controlar y mejorar. De allí la importancia de los datos cuantitativos, de la precisión con respecto a los que son realmente importantes y, de la pericia del personal, en conjunto con las herramientas tecnológicas eficaces para su utilización. | |

**Gobernanza de datos y calidad de la información empresarial**

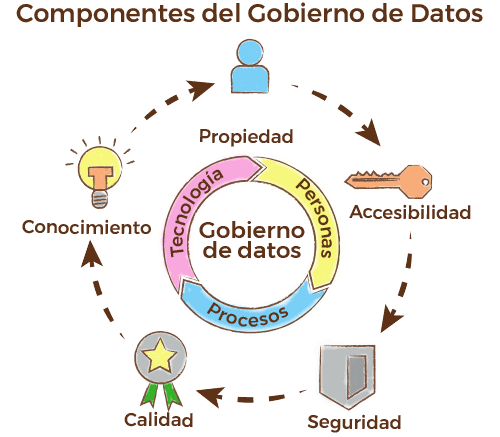
| Cuadro de texto |
| --- |
| Es importante tener presente que, en las empresas, los datos ameritan de una calidad tal, que favorezca la toma de decisiones efectiva y la mejora de los procesos. En esa calidad, se encuentra el tipo de dato, la manera de utilizarlo, de procesarlo, de almacenarlo y de la protección de los datos. A ello corresponde la gobernanza o el gobierno de datos en la empresa, cuyo fin principal es evitar el fraude en la gestión de los datos, en el contexto interno o externo del sistema empresarial, y la predicción de eventos a través del manejo estadístico de la tendencia de datos, protegiendo al sistema de problemas internos que coloquen en riesgo la productividad, la calidad y la competitividad. |

| **Tipo de recurso** | Carrusel de tarjetas | |
| --- | --- | --- |
| **Introducción** | Es así como la gobernanza de los datos define: | |
| La forma y fuentes de obtención de los datos. | | fuente de datos icono gratis |
| La manera de generar, procesar y almacenar los datos en los procesos. | | procesamiento de datos icono gratis |
| Los datos claves. | | Contraseña de clave, acceso a la información. Elemento de diseño de página.Icono lineal vector aislado en fondo blanco. |
| Los datos innecesarios que deben ser depurados. | | Elimine el concepto de eliminación de datos y mueva los archivos innecesarios al vector de ilustración de la papelera |
| Los datos circunstanciales que pueden obtenerse en momentos coyunturales de la toma de decisiones. | | big data icono gratis |

| Cuadro de texto |
| --- |
| Es por esto que el direccionamiento estratégico, el recurso humano, los procesos y las TICs, son los elementos claves de la gobernanza de los datos en la empresa.  Se denota como gobierno, porque la gestión de los datos y la manera de administrar se somete a políticas propias de la empresa. Son varios los aspectos por los cuales no todas las empresas administran este concepto: por no ser la gestión de los datos un elemento estratégico de la gestión; porque los datos que se manejan no son cuantiosos ni complejos; porque la confidencialidad no forma parte de la política empresarial; porque no se poseen los sistemas de información adecuados para el resguardo de los datos. A pesar de todo esto, está demostrada la importancia de la gestión de datos como elemento de productividad, calidad y competitividad empresarial.  En la Figura 1 se presentan los componentes del gobierno o gobernanza de datos. |

**Figura 1**

*Componentes del gobierno o gobernanza de datos*



| Cuadro de texto |
| --- |
| Aunque la gobernanza de los datos sea una decisión estratégica de la organización, para muchas instituciones es, incluso, un mandato de algunas regulaciones nacionales, especialmente cuando las empresas deben dar cuenta de su gestión a entes gubernamentales, o empresas exportadoras e importadoras.  Existen empresas dedicadas a la gobernanza de los datos de grandes empresas, es decir, algunas empresas tercerizan (encargan la gestión) a otras empresas especializadas en sistemas de información, con personal experto multidisciplinario que conoce el comportamiento del sector empresarial y que establece contratos de confidencialidad de datos. Esto se realiza, especialmente, para el uso de datos masivos como, por ejemplo, en empresas multinacionales o con alta diversificación de procesos.  Las empresas que prestan el servicio de gobierno de datos, hacen uso de la inteligencia artificial (*machine learning*) soportada en *software* altamente especializados. Se basan en el modelo de *Business Analytics* (BA) o Analítica de negocios, y ofrecen soluciones confiables e innovadoras en el manejo de los datos.  Para ampliar esta información, diríjase al material complementario y lea los documentos:   * **Gobierno de datos, un potenciador de los sistemas de gestión de la calidad.** * **Modelo de gobernanza de la infraestructura de datos, para el desarrollo de tecnologías emergentes.** |

**2. Transformación de datos para uso empresarial**

| Cuadro de texto |
| --- |
| el hombre de negocios presiona el icono de conformidad en la pantalla virtual. El hombre tocó un cartel del portapapeles con marca de verificación. Empresa, estrategia, negocio, tecnología, finanzas, concepto de trabajo. Cumplimiento de normas: gráfico con iconos.  Los datos puros no son de utilidad para servirse de ellos en la toma de decisiones. No obstante, pueden entrar al sistema empresarial desde fuentes externas y generarse, a su vez, como resultado del funcionamiento de los procesos del sistema. Sin embargo, si el fin es la toma de decisiones, a partir del manejo de indicadores de gestión, entonces, los datos deben ser transformados en información. Este tema está dedicado a la exposición de la transformación de datos para uso empresarial, en donde juegan un rol esencial las personas, las tecnologías, los procedimientos de gestión de los datos, y el establecimiento de la forma en la cual la empresa amerita organizar los datos para su uso.  Primero, es necesario ampliar la información sobre la gestión de los datos; por este motivo, lo invitamos a ver el siguiente video. |

| **Tipo de recurso** | Video motion | | | |
| --- | --- | --- | --- | --- |
| **NOTA** |  | | | |
| **Título** | Gestión de los datos | | | |
| **Escena** | **Imagen** | **Sonido** | **Narración** | **Texto** |
| **1** | Concepto de gestión de proyectos con hombre de negocios azul |  | La gestión de los datos en la empresa, posee dos elementos claves: considerar los datos como activos intangibles que aportan un valor estratégico al desarrollo de los procesos, y contribuir con la optimización de los datos disponibles en el marco de las normativas, políticas, necesidades o direccionamiento organizacional. | Gestión de los datos  Activos intangibles  Optimización de los datos |
| **2** | Ilustración del concepto de método scrum |  | La gestión idónea de los datos debería contribuir con las personas en el desarrollo de los procesos, obteniendo las metas de desempeño esperadas.  Una gestión idónea corresponde a la obtención y transformación de datos de calidad, en información de calidad, es decir, la información realmente necesaria. Asimismo, la gestión asertiva y oportuna de los datos, previene la aparición de eventos adversos a los objetivos organizacionales. | Desarrollo de los procesos  Gestión idónea  Gestión asertiva y oportuna de los datos |
| **3** | Gestión de kpi y tareas. optimización del flujo de trabajo |  | En torno a la gestión de los datos, luego de la obtención, el proceso clave es el de transformación. Esta transformación hace referencia a la conversión de datos simples en datos más complejos, o de los datos puros, en información conclusiva de hallazgos en torno a un fenómeno analizado. En este sentido, la “transformación”, implica un cambio que sufren los datos originales. Dichos cambios son coherentes y consistentes con la “forma” de los datos que amerita la empresa.  Por ejemplo, un grupo de datos puede pasar a conformar el contenido de variables que permitan calcular indicadores de gestión; la recolección de un conjunto de datos con respecto a las cantidades producidas en una línea de producción. Cada día, durante un mes continuo. | Transformación  Conversión  Cambio que sufren los datos originales |
| **4** | Hombre de negocios haciendo presentación con sus colegas y estrategia de negocio efecto de capa digital en la oficina como concepto.Grupo de analistas trabajando en gráficos. |  | El promedio de producción del mes, con el promedio de producción del mes siguiente, permite el cálculo de un índice de desempeño de la producción, llegando a ser el dato simple de las cantidades producidas, ahora un índice de desempeño.  Otro ejemplo es la transformación de datos simples en conjuntos de datos, en rangos de datos, permitiendo conocer en qué rangos se agrupa la mayor cantidad de datos recolectados, apuntando a la tendencia en el comportamiento de los datos. | Índice de desempeño de producción  Conjuntos de datos  Rangos de datos |
| **5** | Médico visitante paciente gordo |  | Asimismo, la transformación de los datos se puede evidenciar en cómo, con un dato, se puede llegar a concluir en torno a una variable. Por ejemplo, si se quiere calcular el índice de masa corporal (IMC) de una persona, se requiere conocer el peso, la estatura, y la circunferencia abdominal.  A través de una fórmula que agrupa los tres datos de medida de la persona, se logra transformar esos tres datos en un dato global, el cual, incluso, puede llegar a generar un IMC promedio de una población total en estudio. | Índice de masa corporal (IMC) |
|  | Diseño de vectores planos, análisis estadístico y de datos para el concepto de inversión de finanzas empresariales con un equipo de personas de negocios que trabaja en el tablero de gráficos de monitoreo |  | En fin, transformar un dato, amerita obtenerlo y generar datos más complejos, sin que el dato sea alterado. Por esto, es fundamental conocer qué datos se requieren, para cuáles fines y cuál va a ser el medio de transformación analítica del dato.  En cuanto a este punto, la tecnología es la aliada fundamental para el manejo matemático y estadístico de los datos, así como el análisis de equipos de expertos, para el caso de datos cualitativos. Se concluye que la transformación de los datos es un proceso formal de la analítica de los datos empresariales. | Qué dato se requiere  Cuál es su fin  Medio de transpormación |
| **Nombre del archivo** | 134104\_v2 | | |  |

**2.1 Herramientas de gestión de datos**

| Cuadro de texto |
| --- |
| concepto de gestión de datos, concepto de tecnología de la información  Por otra parte, las herramientas de gestión de datos son el apoyo tecnológico que utilizan las empresas para obtener, transformar y distribuir los datos para el funcionamiento de los procesos. Esto se basa en el proceso de *Data Management* (DM) o gerencia de datos empresariales.  En la actualidad, con los avances tecnológicos y la evolución empresarial, resulta prácticamente imposible realizar la gestión de los datos sin el apoyo de herramientas tecnológicas, para hacer de los datos diversos, información unificada y organizada. Para esto, se deben cubrir elementos de recursos y capacidades empresariales de alfabetización tecnológica del personal, adquisición de *software*, o tercerización del servicio de gestión de datos.  Las herramientas de gestión de datos, no solo cumplen un rol tecnológico, sino que son un activo estratégico de la analítica de datos en la empresa. Todo lo que tenga que ver con las herramientas de gestión de datos, es considerado una inversión y no un gasto, es decir, que aportan valor a los resultados empresariales. |

| **Tipo de recurso** | Infografía estática |
| --- | --- |
| **Texto introductorio** | A continuación, conozcamos los objetivos de estas herramientas: |
|  | |
| **Código de la imagen** | 134101\_i14 |

| Cuadro de texto |
| --- |
| Para evidenciar cómo, en la actualidad, la gestión asertiva de los datos influye de manera positiva en la inteligencia del negocio, diríjase al artículo **La inteligencia de negocios: herramienta clave para el uso de la información y la toma de decisiones empresariales,** dispuesto en el material complementario. |

| Cuadro de texto |
| --- |
| Es importante tener en cuenta que las herramientas de gestión de datos son activos empresariales, sean controlados por la misma empresa o por organizaciones que prestan el servicio. Estas herramientas cuentan con unas características, las cuales se define a continuación: |

| **Tipo de recurso** | Slider Presentación | |
| --- | --- | --- |
| **Apropiación de la herramienta**  Depende del tipo de empresa y debe ajustarse al tamaño de la empresa, al tipo de procesos, al tipo de informes a generar, y a las capacidades de alfabetización tecnológica del personal de la empresa. Pueden incluso ser exclusivas de la empresa, o ser sistemas de información comercial, obtenidos bajo licencia.  Debe mantenerse el lema que dice: primero es la estrategia y luego la estructura; es decir, que si se amerita de una herramienta X, la empresa debe hacer las adaptaciones necesarias a sus procesos, para poder obtener los resultados deseados con el uso de la herramienta de gestión de los datos. | | Enterprise client está activando tres iconos de servicios administrados en una interfaz de control virtual con botones hexagonales. Concepto de negocio y tecnología de la información para administración externa de TI. Copiar espacio. |
| **Capacidad de la herramienta**  Debe ajustarse a la cantidad de datos a obtener, procesar (transformar), al tipo de informes a generar y a la infraestructura tecnológica de la empresa. La cantidad de datos a gestionar, masivos o no, determina la capacidad tecnológica de la herramienta. | | Análisis digital, análisis de datos masivos, ciencia de datos, investigación de mercado, ilustración de banner de vector plano de aplicación con iconos |
| **Tecnificación de la herramienta**  El creador de *software* incorpora, de acuerdo con la calidad y capacidad de la herramienta, la mayor cantidad y apropiación de algoritmos matemáticos para el procesamiento de datos, considerable capacidad de almacenamiento, y mayores opciones de recursos de análisis.  A mayor complejidad de procesos, mayor complejidad de las herramientas de gestión de datos. | | Científicos de datos. Business Intelligence, concepto de desarrollo de software. Programadora mujer usando laptop analizando información diversa, grafo financiero en pantalla futurista de interfaz virtual |
| **Automatización de los procesos**  Las herramientas de gestión de datos, buscan la automatización de los procesos de analítica de datos, siendo la mano derecha o colaboración esencial de los Administradores de Bases de Datos (ABD) en las empresas.  Estas herramientas permiten pasar de procesos empíricos de gestión de datos, a procesos automatizados que elevan el valor de la toma de decisiones, disminuyendo los errores y dificultades en el manejo manual de los datos. | | Analista trabajando con Business Analytics y Data Management System en el equipo para realizar informes con KPI y métricas conectadas a la base de datos. Estrategia corporativa para finanzas, operaciones, ventas, marketing |
| **Posibilidad de automatización continua**  La tecnología provee a los desarrolladores de herramientas de gestión de datos, su mejoramiento continuo, actualización, cambios, e incluso rediseño, en función de las necesidades de la empresa. Asimismo, los *software* comerciales pueden ser adaptados a las necesidades particulares de los procesos de las empresas. Uno de los fines de la automatización continua, es agilizar las respuestas de las herramientas, a las necesidades de obtención de reportes en la empresa. | | Double exposure of professional businessman hold Database table with business city background |

| Cuadro de texto |
| --- |
| Sistema informático que utiliza para servidores en la nube con sol. Concepción de documentos digitales y mensajes que utilicen para la moderna tecnología de pc conectada. Ilustración vectorial.  Las herramientas de gestión de datos han variado, con el transcurrir del tiempo, al ritmo de la evolución tecnológica. Así, como se conoce a Excel, una herramienta básica de gestión de datos cuantitativos, la tendencia tecnológica en la administración de los datos, ha generado herramientas como la “la nube” o reservorio “no físico” (desde 1990); es decir, no se encuentra alojado en el disco duro del computador, y funciona como almacenador de datos masivos (Solved Magazine, 2020). Esta herramienta es ofrecida por la Internet, y servidores como Google, la maneja en una nube conocida como Drive, que permite a las personas el almacenamiento gratuito hasta una cierta capacidad de datos. Mundialmente, el Google Drive es reconocido como la herramienta que reemplazó a Google Docs, dadas sus mayores y mejores utilidades y capacidades de almacenamiento y transmisión y modificación de la información, en grupos de trabajo. |

**2.2 Análisis exploratorio de datos y alistamiento de los *sets* de datos**

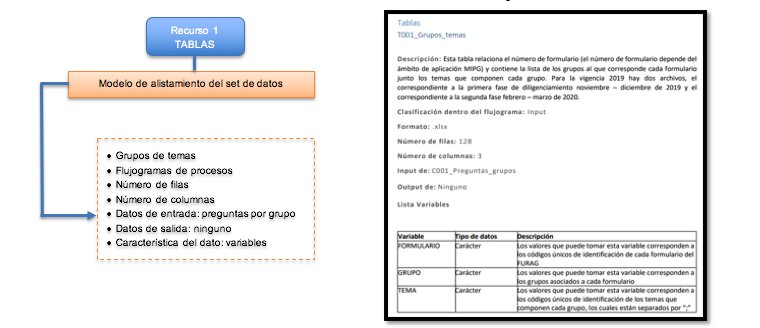
| Cuadro de texto |
| --- |
| Hablemos primero del análisis exploratorio de datos que es un método de selección de entrada de datos al sistema organizacional o al proceso de analítica de datos y que tiene por objetivo la selección de los datos únicos, oportunos y apropiados, para las necesidades de la empresa. El detalle a tener en cuenta es que, este análisis exploratorio de datos se realiza con base en las herramientas de gestión de datos, las cuales fueron creadas a partir de la matemática y la estadística y poseen la capacidad de depurar los datos con técnicas algorítmicas de las cuales derivan los datos de interés para la analítica de los datos. En consecuencia, se obtiene del análisis exploratorio de datos, un resumen o conjuntos de datos relevantes para dicha analítica. informe de presentación con vidrio magnificado  Para conocer todos los elementos técnicos del análisis exploratorio de datos, desde la base de sus creadores principales en el contexto mundial, diríjase al artículo **Análisis exploratorio de datos,** el cual se encuentra disponible en el material complementario. |

| **Tipo de recurso** | | Pestañas o tabs Verticales |
| --- | --- | --- |
| **Introducción** | | Para continuar con el alistamiento y, para su garantizar su comprensión, se dará respuesta a las siguientes preguntas: |
| Crecimiento comercial y económico en la red de negocios global, Análisis de datos de finanzas y banca, Servicio al cliente, Conexión de tecnología y datos, Trabajo en equipo, Estrategia de negocio y marketing digital. | | |
| **¿Qué es el alistamiento de *set* de datos?** | El alistamiento es un término técnico usado en tecnología para el manejo o gestión de los datos; un “*set*” de datos es análogo a un grupo de datos. El alistamiento de *set* de datos corresponde a la preparación o armado de una colección de datos para su uso, a través de las herramientas tecnológicas, es decir, es el proceso de organización previo a la distribución de los datos para su uso en el sistema empresarial, sea para el inicio de los procesos, o para mostrar el desempeño final de los mismos y pasar a la retroalimentación del sistema y a la nueva toma de decisiones. | |
| **¿Qué implica alistar en *sets*?** | “Alistar en *sets*”, implica preparar los datos en agrupamientos más complejos como variables, conjuntos o rangos de datos, los cuales, al estar renombrados y agrupados con códigos por tipos de datos, permiten un mejor registro y utilización. El alistamiento se realiza a través de las herramientas de gestión de datos por parte de los programadores, razón por la cual es totalmente necesaria la multidisciplinariedad en la administración empresarial, en el caso que se quiera hacer de los datos un activo real de la gestión empresarial. | |
| **¿Qué sucede cuando se poseen los datos?** | Una vez se poseen los datos, sea guardados en el computador para su procesamiento con herramientas básicas como Excel o para *software* especializados, los mismos deben reestructurarse, agruparse, conglomerarse o mostrarse en recursos visuales o gráficos de mayor utilidad; tal es el caso de tablas, gráficos, u hojas de documentos, bajo un orden específico, que se determina a partir del tipo de dato y de la funcionalidad para el análisis. | |

| Cuadro de texto |
| --- |
| En Colombia, una institución que se puede tomar, como ejemplo, en cuanto a la preparación y uso de set de datos por la complejidad de los datos que maneja, es el Departamento Administrativo de la Gestión Pública, ente que ofrece, al público en general, su modelo de alistamiento y distribución de datos en el caso particular de la medición del desempeño institucional, bajo el esquema del modelo figurativo FURAG, siglas del “Formulario Único de Reportes y Avances de Gestión”, constituido como registro administrativo y creado con el propósito de recolectar datos sobre el avance en la implementación de las políticas de gestión y desempeño institucional, que hacen parte del Modelo Integrado de Planeación y Gestión (MIPG) de la gestión pública nacional. Si se revisa este modelo, se tiene un alistamiento de *set* de datos estratégico construido como se muestra en la Figura 2. |

**Figura 2**

*Modelo de alistamiento de set de datos*



Nota. Tomado y adaptado de Departamento Administrativo de la Gestión Pública (2020).

| Cuadro de texto |
| --- |
| En cualquier organización, se alistan los datos para su procesamiento; tal es el caso de alistamiento de datos para producción, para almacenaje, o para procesos de consultoría como, por ejemplo, el correspondiente a un proceso de Consultoría en el Plan de Ordenamiento Territorial del Departamento del Huila (Consorcio POT Huila, 2019), Colombia, el cual puede encontrar completo en el material complementario.  En la Figura 3 se muestra un ejemplo de cómo el analista, modeló su esquema de alistamiento de datos, y luego de la obtención de los datos pasó a manejarlo a través de herramientas de gestión de datos. |

**Figura 3**

*Modelo de alistamiento de set de datos para el proceso de consultoría*



Nota. Tomado de Consorcio POT Huila – Informe etapa de alistamiento (2019).

| Cuadro de texto |
| --- |
| Con estos ejemplos se quiere mostrar que el alistamiento de datos no posee un modelo estandarizado, sino que se prepara en función de las necesidades del analista de datos y de las herramientas tecnológicas disponibles para tal fin. |

**3. Las bases de datos**

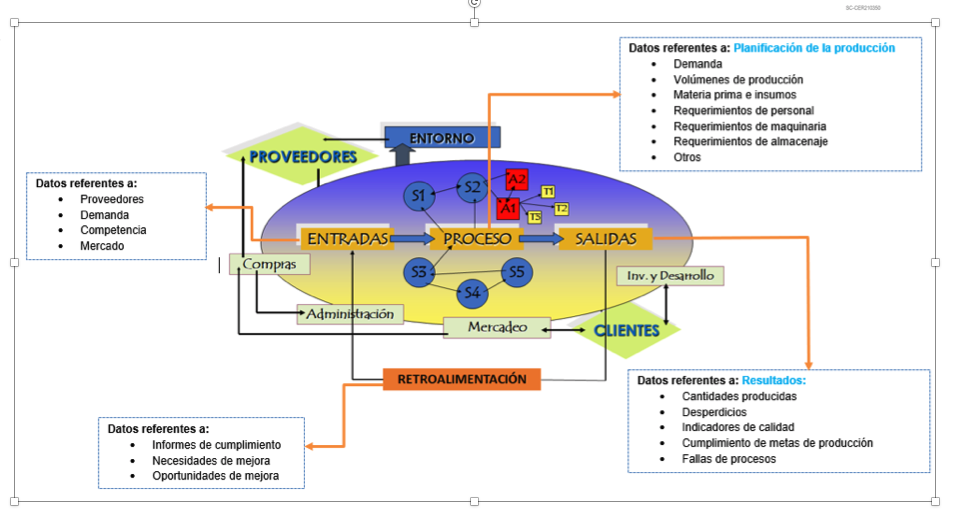
| Cuadro de texto |
| --- |
| Diseño de gráficos de información de concepto de servidor de base de datos y de red, vector limpio  Este tema se dedica a la exposición de la conceptualización, características y tipos de base de datos, como elemento del proceso de comunicación e información de los sistemas empresariales. Asimismo, encontrará al detalle, el desarrollo de los procesos ETL (Proceso de extracción, transformación y carga) en los que se sustenta la gestión de los datos, desde las bases de datos. Con estos elementos, podrá comprender la utilidad de la tenencia, actualización, mantenimiento y resguardo de las bases de datos en las empresas, las cuales son un reservorio histórico y actual de datos de personas, procesos y recursos, utilizados para poner en funcionamiento el sistema y que además muestran el desempeño paulatino del sistema empresarial. |

**3.1 Tipos de bases de datos**

| Cuadro de texto |
| --- |
| Hasta este punto se ha estudiado que los datos poseen diferentes fuentes, tipos y maneras de procesarlos; además, diversos datos satisfacen distintas necesidades para los procesos empresariales. Lo importante es retomar que, en todo el contexto del sistema empresarial, se ameritan datos, como se muestra en la Figura 4. |

**Figura 4**

*Necesidades de entradas (inputs) y generación de datos sistema empresarial*



| Cuadro de texto |
| --- |
| Como se observa en la Figura 4, si la empresa no posee herramientas tecnológicas que le permitan almacenar los datos, sería muy difícil hacer análisis de datos históricos, análisis en tiempo presente y proyecciones futuras, de comportamiento del sistema. Es por esta razón que las bases de datos son reservorios o almacenes de datos que pueden ser compartidas en los sistemas empresariales e incluso, con sistemas de otras empresas, siempre que sean compatibles. |

| **Tipo de recurso** | Acordeón tipo 1 |
| --- | --- |
| **Introducción** | Las bases de datos se pueden clasificar según su relación con el sistema organizacional; así que pueden ser: |
| Illustration of service oriented layer architecture and how it connects to various external, internal systems | |
| **Bases de datos externas**  Si se nota el elemento de las entradas, datos como los de la competencia, del tamaño y ubicación del mercado, o de la red de proveedores de un sector empresarial, pueden encontrarse en bases de datos externas. No todo el tiempo se poseen datos organizados como set de datos fácilmente utilizables por la empresa, sino que se consiguen datos aislados desde diferentes fuentes que deben ser alistados como *sets* de datos.  Por el contrario, otras fuentes externas manejan bases de datos debidamente organizadas; tal es el caso de los entes gubernamentales, en los que es política de la gobernanza de datos, mantener bases de datos de uso público, como lo determinan los gobiernos en sus políticas mundiales. Por ejemplo, en Colombia, el Departamento Administrativo Nacional de Estadística (DANE), debe construir y publicar bases de datos en términos económicos, demográficos, de salud, entre otros, exigidos por normativa internacional. | |
| **Bases de datos internas**  Retomando el esquema de la figura 4, se tienen datos internos de desempeño como las cantidades producidas por tipo de producto, datos de desperdicio, satisfacción del cliente, fallas de procesos, entre otros, que son almacenados en bases de datos, particularmente en Excel (a menos que se posean *software* especializados), y que son utilizados para la preparación de los informes de gestión.  Las bases de datos internas, normalmente no poseen una estructura estándar, debido a que cada proceso, dentro de la empresa, maneja sus propias variables y tipos de datos. No obstante, al momento de preparar los indicadores de desempeño globales, es importante tener coordinación en el tipo de dato que se agrupa del total de los procesos. | |

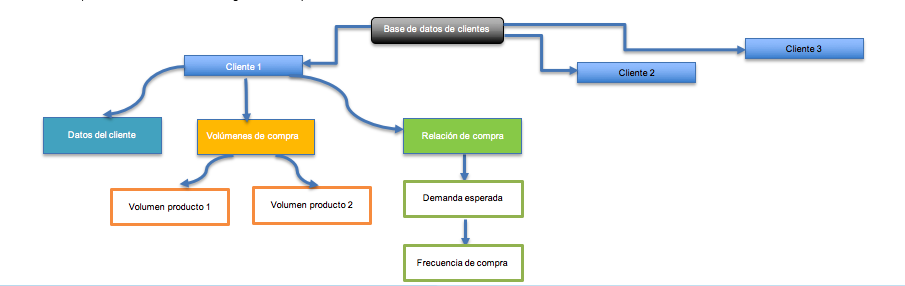
| Cuadro de texto color |
| --- |
| La importancia que recae en las bases de datos en la empresa, sean internas o externas, es la de tener los datos oportunos y veraces, en el momento y lugar que se requiere, para agilizar los procesos, incrementando la productividad empresarial.  Ahora, desde el punto de vista del tipo de dato, específicamente de su complejidad y de la relación entre varios datos, se hace necesario tener en cuenta algunos tipos de bases de datos que se han ido generando en la medida que crecen los avances tecnológicos y la satisfacción de las necesidades de quien usa los datos. Conozcamos algunas de esas bases de datos. |

**Bases de datos jerárquicas**

| Cuadro de texto |
| --- |
| Estas bases de datos se construyen cuando se manejan datos dependientes o relacionados, es decir, datos cuyo resultado está relacionado con un dato previo, y que luego genera otro dato consecuente. Por ejemplo, los datos de los padres generan información con respecto a los hijos y de estos, hacia los hijos de los hijos (nietos); de esta manera, se puede tratar una base de datos de tipo generacional.  Ahora, en la Figura 5 se observa un ejemplo relacionado con la empresa. |

**Figura 5**

*Ejemplo de base de datos jerárquica*



| Cuadro de texto |
| --- |
| La característica fundamental de este tipo de bases de datos, es que en físico se ve como un árbol con ramas y sub-ramas, en las cuales se observan las jerarquías desde lo principal a lo que, en escalada, va tomando su lugar en dichos ramales. Entonces, en estas bases de datos, se almacena la información en forma de registros con jerarquías, o el orden de aparición, o el orden de mayor a menor importancia de los datos, o el orden de aparición de los datos en el fenómeno estudiado.  En este tipo de bases de datos, técnicamente, cada jerarquía se denota como “nodo”. En el árbol habrá un “nodo raíz” y, a partir de este nodo, se irán creando otros nodos, todos interrelacionados. |

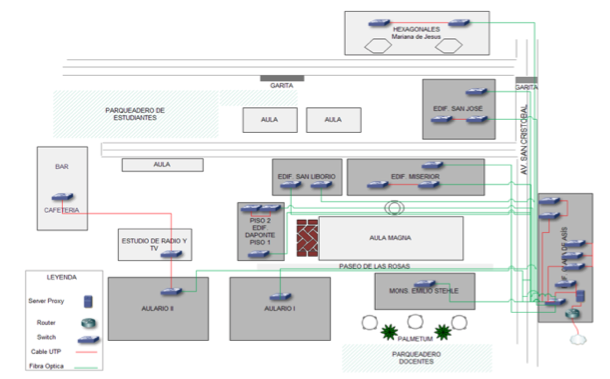
| **Tipo de recurso** | | Pestañas o tabs Verticales |
| --- | --- | --- |
| **Introducción** | | En estas bases de datos, deben tenerse en cuenta ciertos aspectos como: |
| Diseño de iconos de color vectorial del sistema de archivos en clúster, símbolo de datos gigantes, signo de inteligencia empresarial, alojamiento web e iluminación de stock del centro de datos, conceptos de cubos de datos OLAP | | |
| **Datos de diversas fuentes** | **Para alimentar la base de datos, se ameritan los datos de diversas fuentes.** Por ejemplo, harán falta datos de fuentes externas, o datos de diferentes unidades organizativas. En este caso, cuando se poseen sistemas de información dentro de la empresa, es fundamental mantener al día, la carga de los datos. | |
| **Datos independientes** | **Los datos son independientes.** Pese a que se encuentran relacionados, corresponden a distintos datos, o son datos de diversas variables; no obstante, son variables interrelacionadas. Es decir, el comportamiento de un dato no depende obligatoriamente del comportamiento del otro dato. | |
| **Origen del dato** | **Los datos poseen su propio origen.** Con esto se quiere resaltar que la naturaleza del dato no puede, ni debe modificarse. | |
| **Naturaleza del dato** | **Se requiere conocimiento de las personas o procesos, en cuanto a la naturaleza del dato.** El dato amerita conocerse en toda su extensión; para qué sirve, con qué se relaciona, cómo se obtiene, cómo debe almacenarse, cómo ese dato aporta en la toma de decisiones de la empresa, cómo ese dato, afecta a los demás datos de la jerarquía. | |
| **Relación de los nodos** | **La relación de los nodos debe mantenerse.** Dependiendo de la cuantía y cantidad de niveles jerárquicos, puede correrse el riesgo que, en el análisis, se vaya perdiendo el relacionamiento. Es por esto que las personas, procesos y recursos involucrados en la gestión de la jerarquía de los datos, son fundamentales. | |

**Bases de datos de red**

| Cuadro de texto |
| --- |
| En este caso, a diferencia de la jerarquía, en la cual un dato depende de solo un nivel jerárquico anterior, en las bases de datos en red, los datos pueden tener más de un relacionamiento, y no interesa la jerarquía, sino los nodos de relación. Por ejemplo, en un problema de producción, pueden estar involucrados los datos de desempeño del talento humano, o los datos de la cantidad de proveedores y tiempos de entrega, o los datos de fallas de la maquinaria. Por esto, al estar un nodo relacionado con muchos más, se hacen bases de datos más complejas.  La forma de este tipo de base de datos se puede observar a continuación, con respecto a la instalación de un sistema interconectado de datos en una universidad. En la Figura 6 se observa el esquema gráfico de ubicación de interés de la red de datos, en función de la infraestructura académica y administrativa. |

**Figura 6**

*Esquema gráfico de ubicación de interés de la red de datos en función de la infraestructura académica y administrativa*

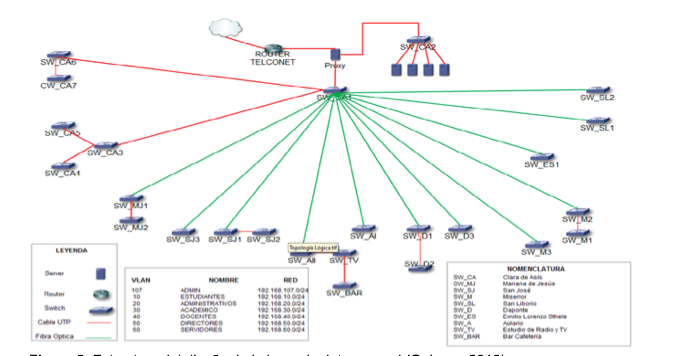


Nota. Tomado de Galarza (2018).

| Cuadro de texto |
| --- |
| Mientras que en la siguiente Figura 7, se muestra el diseño de nodos de una base de datos de la red, con base en el origen de los datos. |

**Figura 7**

*Estructura del diseño de la base de datos en red*



Nota. Tomado de Galarza (2018).

**Bases de datos relacionales**

| Cuadro de texto |
| --- |
| concepto de ilustración vectorial de la tabla de datos relacional  Las bases de datos relacionadas son aquellas bases de datos de diferentes fuentes, que al ser coherentes y consistentes, pueden ser fácilmente relacionadas, comparadas y contrastadas; es decir, son bases de datos independientes, pero que poseen características similares en su diseño, y se hace fácil, para el analista, utilizarlas como base de comparación, e incluso para alimentar de la mejor manera, la base de datos original. Por ejemplo, al interior de la empresa, pueden ser base de datos relacionales las de producción de la línea 1 de ensamble de un producto, con los de la línea 2; o en el caso de producción de panes, la base de datos de costos de producción (estructura de costos), de los panes del tipo 1, con los de los panes de tipo 2. En el caso que la productividad de un tipo de pan sea mejor que la del otro tipo, pueden compararse datos y detectarse variables claves que pueden aportar a la mejora. |

**Bases de datos multidimensionales (MDB)**

| Cuadro de texto |
| --- |
| Concepto de diminuta persona de procesamiento analítico o multidimensional en línea o OLAP. Análisis eficaz de sistemas de datos empresariales con ilustraciones vectoriales de abordaje complejas de bases de datos. Gestión de herramientas de informes de empresa  Por otro lado, las bases de datos multidimensionales son las bases de datos más complejas por la cantidad de datos que relacionan, la cantidad de tipo de datos, los atributos de los datos, entre otros aspectos. La construcción la realizan expertos, pero una vez diseñadas en la empresa, es posible su uso para el desarrollo de los procesos. Se visualizan para los diseñadores como “cubo de datos” y, en programación informática, los datos se almacenan interconectados.  Para ampliar la información sobre su diseño y utilidad, lo invitamos a ver el video **Base de datos multimensionales,** el cual se encuentra en el material complementario. |

| **Tipo de recurso** | Tarjetas Avatar | |
| --- | --- | --- |
| **Introducción** | Teniendo en cuenta lo anterior, hablemos sobre la clasificación de las bases de datos según su necesidad de modificación y adaptación.Desde el punto de vista tecnológico, las bases de datos poseen tipologías particulares, las cuales tienen que ver con su flexibilidad de modificación con base en las necesidades de los datos y de los modelos de gestión de la empresa, e incluso, con base en los avances tecnológicos. Estas clasificaciones se refieren a: | |
| **Bases de datos dinámicas**  Los datos pueden actualizarse o hasta modificarse; por ejemplo, lo que ocurre durante todo un día de producción en la empresa, en las entregas a los clientes; una de las más fáciles de entender, se refiere a las de los indicadores económicos mundiales, cuyos datos pueden cambiar en cuestión de minutos y tal es el caso del precio del dólar. | | Ingeniero de fábrica de automóviles en uniforme de trabajo usando computadora portátil con software de hoja de cálculo. Trabajar con software en la planta de fabricación industrial para automóviles dedicada a la producción de vehículos. |
| **Bases de datos estáticas**  Se trata de “bases de datos de consulta”. Los datos son originales, emitidos por fuentes externas a la empresa o al usuario. Se usa tal y como está diseñada desde su procedencia. | | Tecnología de inteligencia de negocios y análisis de big data medios mixtos que muestran el concepto de informe de información futurista usando software de computadora para analizar el asesoramiento estratégico de inversión para la toma de decisiones. |

| Cuadro de texto color |
| --- |
| Es importante denotar que aunque, como aprendiz, no posea conocimiento técnico en el diseño de bases de datos bajo programación de sistemas, las bases de datos de sistemas de información adquiridos como *software*, son diseñadas por los programadores (usualmente ingenieros informáticos o de sistemas), con algoritmos lógicos, recursos de buscadores, reglas informáticas, entre otros elementos. No es la intención solo aprender este oficio, sino comprender la utilidad del uso de las bases de datos en la empresa, y de la inmersión del aprendiz en la obtención, análisis y difusión de los datos a partir del uso de una base de datos. |

**3.2 Procesos ETL para el manejo de bases de datos**

| Cuadro de texto |
| --- |
| Los procesos ETL en sus siglas en inglés, sintetizan las actividades de extracción (E, *extract*), transformación (T, *transform*) y alistamiento, carga o almacenamiento (L, *load*), de los datos. Inicia en el momento en que un analista de datos o los involucrados en el proceso de analítica de datos, obtienen (extraen) los datos desde sus diversas fuentes, los agrupan y transforman en función de las necesidades de análisis, y los alistan, almacenan o disponen para su uso, por las partes interesadas.Nube o en los procesos locales de ETL y ELT. Extraer datos de transformación y carga en Base de datos y Almacén de datos. Procesamiento resistente y distribuido de datos estructurados, csv, json, tabla.  La importancia de darle un nombre a estos tres procesos, es que el personal de la empresa, en los niveles estratégico, táctico y operativo, comprenda que la gestión de los datos no es aleatoria, sino que amerita procesos formales y darle a los datos, su debida importancia en el logro del desempeño de la organización. |

| **Tipo de recurso** | Infografía interactiva Punto caliente | |
| --- | --- | --- |
| **Texto introductorio** | Los tres puntos claves en los procesos ETL son: | |
|  | | |
| **Código de la imagen** | 134101\_i37 | |
| **Punto caliente 1** | * Tener claridad del tipo de dato necesario. * Obtener datos de fuentes de datos relevantes, veraces y permanentemente disponibles. * Organizar los datos en función de las necesidades de análisis. * Conocer el modelo de base de datos requerida por la empresa. | EXTRACCIÓN |
| **Punto caliente 2** | * Obtener datos consistentes y coherentes a partir de la depuración. * Convertir, si es necesario, datos en otras unidades de análisis, de acuerdo con las necesidades de la empresa (compatibilidad de datos). * Realizar nuevos procesos de depuración de datos irrelevantes. * Ordenar y filtrar datos. * Muchas veces es necesaria la traducción de datos en un lenguaje comprensible al interior de la empresa. | TRANSFORMACIÓN |
| **Punto caliente 3** | * Definir el proceso de carga de los datos, la cual puede ser directa en función de la base de datos de origen, si es de fuentes externas o de otros procesos, o carga manual en la realidad diaria de los procesos, o carga mecánica programada con los equipos de producción. * Con el apoyo de expertos en informática, carga de datos complejos que requieren ser cargados bajo lenguajes de programación. | CARGA |

| Cuadro de texto |
| --- |
| Comprender esta nomenclatura para la gestión de los datos en la analítica de datos que amerita la inteligencia empresarial, implica que cobran importancia los sistemas de información en la empresa, surgiendo la necesidad de optar por Sistemas de Gestión de Bases de Datos (SGBD), cuya mayor formalidad y tecnificación, proveerá mayor calidad a la obtención de información para la toma de decisiones. |

| Cuadro de texto |
| --- |
| Ha finalizado el estudio de este recurso educativo. Recuerde explorar los demás recursos que se encuentran disponibles en este componente formativo; para ello, diríjase al menú principal, donde encontrará la síntesis, una actividad didáctica para reforzar los conceptos estudiados, material complementario, entre otros. |

**SÍNTESIS**

| **Tipo de recurso** | Síntesis |
| --- | --- |
| Asistencia para la inteligencia empresarial  Síntesis: Procedimiento de analítica de datos | |
| **Introducción** | El siguiente mapa integra los criterios y especificidades de los conocimientos expuestos en el presente componente formativo. |
|  | |

**ACTIVIDAD DIDÁCTICA**

| **Tipo de recurso** | Actividad didáctica. Verdadero y falso | |
| --- | --- | --- |
| Apreciado aprendiz, a continuación encontrará una serie de preguntas que deberá resolver, con el objetivo de evaluar la aprehensión de los conocimientos expuestos en este componente formativo. | | Gestión y desarrollo del CMS, concepto de administración de contenidos de sitios web. Administrador trabajando con base de datos, sistemas de tecnología de la información. Ilustración vectorial plana aislada en fondo blanco |
| 1. La analítica y resguardo de los datos en la empresa, es esencial en cualquier tipo de organización, indistinto de su tamaño y sector productivo. | |  |
| Verdadero (correcto) | | Falso |
| Retroalimentación positiva:  ¡Muy bien! Ha elegido la respuesta correcta.  Retroalimentación negativa:  ¡Incorrecto! Es necesario revisar nuevamente el tema sobre la analítica y el resguardo de los datos en la empresa. | | |
| 2. La analítica de datos es un proceso que amerita solo de las tecnologías. | |  |
| Verdadero | | Falso (correcto) |
| Retroalimentación positiva:  ¡Muy bien! Ha elegido la respuesta correcta.  Retroalimentación negativa:  ¡Incorrecto! Es necesario revisar nuevamente el tema sobre la analítica de datos. | | |
| 3. Los datos en la empresa provienen de distintas fuentes y son de diferentes tipos. | |  |
| Verdadero (correcto) | | Falso |
| Retroalimentación positiva:  ¡Muy bien! Ha elegido la respuesta correcta.  Retroalimentación negativa:  ¡Incorrecto! Es necesario revisar nuevamente el tema sobre tipos de bases de datos. | | |
| 4. La gobernanza de datos implica que la gerencia es quien maneja los datos de la empresa. | |  |
| Verdadero (correcto) | | Falso |
| Retroalimentación positiva:  ¡Muy bien! Ha elegido la respuesta correcta.  Retroalimentación negativa:  ¡Incorrecto! Es necesario revisar nuevamente el tema sobre la analítica y resguardo de los datos en la empresa. | | |
| 5. Los datos, antes de ser utilizados, deben ser transformados. | |  |
| Verdadero (correcto) | | Falso |
| Retroalimentación positiva:  ¡Muy bien! Ha elegido la respuesta correcta.  Retroalimentación negativa:  ¡Incorrecto! Es necesario revisar nuevamente el tema sobre transformación de datos para uso empresarial. | | |
| 6. La tecnología es la principal herramienta de gestión de datos en las empresas. | |  |
| Verdadero (correcto) | | Falso |
| Retroalimentación positiva:  ¡Muy bien! Ha elegido la respuesta correcta.  Retroalimentación negativa:  ¡Incorrecto! Es necesario revisar nuevamente el tema sobre las herramientas de gestión de datos. | | |
| 7. El análisis exploratorio de datos, corresponde a la creación de nuevos datos. | |  |
| Verdadero | | Falso (correcto) |
| Retroalimentación positiva:  ¡Muy bien! Ha elegido la respuesta correcta.  Retroalimentación negativa:  ¡Incorrecto! Es necesario revisar nuevamente el tema sobre el análisis exploratorio de datos. | | |
| 8. El alistamiento de los *sets* de datos, es la primera etapa en la obtención de los datos. | |  |
| Verdadero | | Falso (correcto) |
| Retroalimentación positiva:  ¡Muy bien! Ha elegido la respuesta correcta.  Retroalimentación negativa:  ¡Incorrecto! Es necesario revisar nuevamente el tema sobre el alistamiento de los sets de datos. | | |
| 9. Existen diferentes tipos de bases de datos y cada empresa utiliza la más conveniente de acuerdo con sus procesos. | |  |
| Verdadero (correcto) | | Falso |
| Retroalimentación positiva:  ¡Muy bien! Ha elegido la respuesta correcta.  Retroalimentación negativa:  ¡Incorrecto! Es necesario revisar nuevamente el tema sobre los tipos de bases de datos. | | |
| 10. Los procesos ETL corresponden al diseño tecnológico de las bases de datos. | |  |
| Verdadero | | Falso (correcto) |
| Retroalimentación positiva:  ¡Muy bien! Ha elegido la respuesta correcta.  Retroalimentación negativa:  ¡Incorrecto! Es necesario revisar nuevamente el tema sobre los procesos ETL para el manejo de bases de datos. | | |

Retroalimentación general positiva:

¡Felicitaciones! Ha logrado una óptima aprehensión de los conocimientos relacionados con el procedimiento de analítica de datos.

Retroalimentación general negativa:

¡Inténtelo de nuevo! Lo invitamos a revisar nuevamente el material de estudio para afianzar los conocimientos presentados. ¡Ánimo!

**MATERIAL COMPLEMENTARIO**

| Tipo de recurso | Material complementario | | |
| --- | --- | --- | --- |
| Tema | Referencia APA del material | Tipo | Enlace |
| La analítica y resguardo de los datos en la empresa | Portilla, J. (2017). Gobierno de datos, un potenciador de los sistemas de gestión de calidad. Ministerio de Hacienda y Crédito Público, Colombia. *Revista Signos, 9*(2). p.159-172. | Artículo | <https://www.redalyc.org/journal/5604/560459719010/html/> |
| La analítica y resguardo de los datos en la empresa | Guío, A. (2021). *Modelo de gobernanza de la infraestructura de datos para el desarrollo de tecnologías emergentes.* <https://dapre.presidencia.gov.co/TD/Modelo-Gobernanza-Infraestructura-Datos-desarrollo-tecnologias-emergentes.pdf> | PDF | <https://dapre.presidencia.gov.co/TD/Modelo-Gobernanza-Infraestructura-Datos-desarrollo-tecnologias-emergentes.pdf> |
| Herramientas de gestión de datos | García-Jiménez, A., Aguilar-Morales, N., Hernández-Triano, L. / Lancaster-Díaz, E. (2021). La inteligencia de negocios: herramienta clave para el uso de la información y la toma de decisiones empresariales. *Revista de Investigaciones Universidad del Quindío, 33*(1). p.132-139. <https://ojs.uniquindio.edu.co/ojs/index.php/riuq/article/view/514> | Artículo | <https://ojs.uniquindio.edu.co/ojs/index.php/riuq/article/view/514> |
| Análisis exploratorio de datos y alistamiento de los sets de datos | IBM Cloud Education. (2020). *Análisis exploratorio de datos.* IBM Cloud Education. <https://www.ibm.com/co-es/cloud/learn/exploratory-data-analysis> | Página web | <https://www.ibm.com/co-es/cloud/learn/exploratory-data-analysis> |
| Análisis exploratorio de datos y alistamiento de los sets de datos | Consorcio POT Huila. (2019). *Informe etapa de Alistamiento. Contrato de consultoría no. 1029 de 2019. “Elaborar el plan de ordenamiento territorial del Departamento del Huila”.* <https://www.sirhuila.gov.co/wp-content/uploads/2022-ACTUALIZACIONES/POD/Alistamiento/Informe-Etapa-Alistamiento.pdf> | PDF | <https://www.sirhuila.gov.co/wp-content/uploads/2022-ACTUALIZACIONES/POD/Alistamiento/Informe-Etapa-Alistamiento.pdf> |
| Tipos de bases de datos | Departamento de Computación, UNAN-Managua. (2014). *Bases de datos multidimensionales. UNAN-Managua: Departamento de Computación* (video). YouTube. <https://www.youtube.com/watch?v=CdbWOLcBHNQ> | Video | <https://www.youtube.com/watch?v=CdbWOLcBHNQ> |

**GLOSARIO**

| **Tipo de recurso** | Glosario |
| --- | --- |
| **Alistamiento de datos:** | etapa de la gestión de datos, en la cual se organizan y preparan los datos en la forma y necesidades de la analítica de los datos, por parte de quien amerita los datos para su uso en los procesos empresariales. Es un proceso fundamentado en las herramientas básicas o especializadas de gestión de datos. |
| **Análisis exploratorio:** | es el proceso que emprende el analista de datos para la obtención de la mayor cantidad y calidad de datos a ser utilizados desde diferentes fuentes, en el proceso de analítica de datos. |
| **Analítica de datos:** | proceso de deconstrucción, reconstrucción, análisis, síntesis y transformación de datos, en información necesaria para los procesos de la empresa. |
| **Bases de datos:** | reservorios de datos estandarizados o dinámicos, de fuentes externas o internas a la empresa y que son manejadas, en su mayoría, por herramientas de gestión de datos, es decir, *softwares* especializados. |
| **Calidad de la información:** | es el elemento fundamental en el proceso exploratorio de los datos, la cual se decide en función de la necesidad de datos de la empresa para la toma de decisiones o inicio de los procesos de trabajo. Datos de calidad, obtenidos de fuentes de calidad, y procesados con herramientas de calidad, garantizan un proceso de eficaz de analítica de datos. |
| **Datos:** | es el elemento sometido al proceso de analítica para la obtención de la información que dinamiza los procesos y en el que se fundamenta la toma de decisiones. Pueden provenir de fuentes externas o internas al sistema organizacional. |
| **Extracción de datos:** | es el proceso de seleccionar, obtener y organizar los datos de las bases de datos disponibles o de los procesos de trabajo para la construcción de bases de datos. |
| **Gestión de datos:** | es el proceso de manejo de los datos dentro de la empresa, en el cual se encuentran involucradas las personas, los procesos y las tecnologías. |
| **Gobernanza de datos:** | elemento estratégico de la analítica y gestión de datos de la empresa, que tiene como fin definir las políticas de gestión de datos, y que posee como fin la protección de los datos. |
| **Información:** | resultado del procesamiento organizado y sistematizado de los datos. |
| **Procesamiento de datos:** | denominado analítica de datos, y conformado por los sub-procesos de extracción, transformación, carga y distribución de los datos en la empresa. |

**REFERENCIAS BIBLIOGRÁFICAS:**

| **Tipo de recurso** | Bibliografía |
| --- | --- |
| Consorcio POT Huila. (2019). *Informe etapa de Alistamiento. Contrato de consultoría no. 1029 de 2019 “Elaborar el plan de ordenamiento territorial del Departamento del Huila”.* <https://www.sirhuila.gov.co/wp-content/uploads/2022-ACTUALIZACIONES/POD/Alistamiento/Informe-Etapa-Alistamiento.pdf> | |
| Departamento Administrativo de la Gestión Pública. (2020). *Documento técnico de captura y procesamiento de datos de la Medición del Desempeño Institucional Proceso de Generación de Productos y Servicios para la Gestión Pública.* <https://www.funcionpublica.gov.co/documents/34645357/34703525/Documento_tecnico_captura_procesamiento_datos_mdi.pdf/44b2ea88-5af6-4eb4-56fd-59cd12633e47?t=1594265201214> | |
| Galarza, C. (2018). Diseño e implementación de una red de datos segura para la Pontificia Universidad Católica del Ecuador, Santo Domingo. *Revista Científica Dominio de las Ciencias, 4*(2). p.123-137. <https://www.dominiodelasciencias.com/ojs/index.php/es/article/view/781/901> | |
| Kingatúa, A. (2020). *Las 10 mejores herramientas de gestión de datos para empresas medianas y grandes. Geekaflare*. <https://geekflare.com/es/data-management-software/> | |
| Rodríguez J. (2017). *Los principales softwares de investigación cualitativa.* Fundación para la Investigación Social Avanzada. <https://isdfundacion.org/2017/09/25/software-de-investigacion-social-y-de-mercados-investigacion-cualitativa/> | |
| Rodríguez J. (2017). *Los principales softwares de investigación cuantitativa.* Fundación para la Investigación Social Avanzada. <https://isdfundacion.org/2017/10/10/software-de-investigacion-social-y-de-mercados-investigacion-cuantitativa/> | |
| Solved Magazine. (2020). *The history of cloud computing*. <https://solved.scality.com/solved/the-history-of-cloud-computing/#:~:text=The%20term%20%E2%80%9Ccloud%20computing%E2%80%9D%20itself,in%20academic%20work%20before%20that> | |