**FORMATO PARA EL DESARROLLO DE COMPONENTE FORMATIVO**

|  |  |
| --- | --- |
| PROGRAMA DE FORMACIÓN | Tecnología en Desarrollo e implementación de soluciones para la transformación digital |

|  |  |  |  |
| --- | --- | --- | --- |
| COMPETENCIA | 220501114- Sistematizar datos masivos de acuerdo con métodos de analítica y herramientas tecnológicas. | RESULTADOS DE APRENDIZAJE | 220501114-01 - Preparar datos de según las fuentes de información. |

|  |  |
| --- | --- |
| NÚMERO DEL COMPONENTE FORMATIVO | CF13 |
| NOMBRE DEL COMPONENTE FORMATIVO | Modelamiento, análisis y preparación de datos |
| BREVE DESCRIPCIÓN | El recurso educativo presenta contenidos y técnicas sobre conceptos, teorías y herramientas empleadas en sistematización de datos masivos. Se dan las pautas y contextos sobre los paradigmas más usados para la gestión de información enfocado a la analítica y carga masiva. |
| PALABRAS CLAVE | Analítica, datos, inteligencia de negocio, origen de datos. |

|  |  |
| --- | --- |
| ÁREA OCUPACIONAL | VENTAS Y SERVICIOS |
| IDIOMA | Español |

1. **TABLA DE CONTENIDOS:**

**Introducción**

**1. Modelamiento y gestión de bases de datos**

1.1 Metodologías de diseño y normalización Principios de ACID

1.2 Técnicas de almacenamiento de datos y consultas

1.3 Estándares técnicos

1.4 Ordenamiento de datos, indexación y recuperación

**2. Preparación de datos**

2.1 Entendimiento de la data

2.2 Detección de errores y datos faltantes

2.3 Identificación de variables importantes

2.4 Dataset

**3. La inteligencia de negocios**

3.1 Identificación de las preguntas básicas

3.2 Metodología de integración

3.3 Herramientas de administración

3.4 Técnicas de solución de problemas (modelación de datos)

3.5 Metodologías de análisis (Kimball, Inmon)

3.6 Verificación de valores y escalas

3.7 Procedimientos almacenados y funciones

3.8 Disparadores

**4. Análisis exploratorio de datos**

4.1 Estadística descriptiva y estadística inferencial

4.2 Población y muestra

4.3 Escalas de medida y clasificación de variable

4.4 Técnicas de análisis estadístico

**5. Métodos para hacer análisis exploratorio de datos**

5.1 Datos univariables

5.2 Datos bivariados

5.3 Datos multivariables.

5.4 Reglas de negocio

5.5 Tipo de restricciones

5.6 Programación transaccional

5.7 Programación de estructuras no lineales, desnormalización, series y dataframes

5.8 Álgebra relacional

5.9 SQL

5.10 No SQL

5.11 JSON, BSON y XML

5.12 DDL, DML, DCL

**6. Estructuras y componentes de analítica de datos**

6.1. Bodega de datos

6.2 Tipos estrella

6.3 Copo de nieve

6.4 Constelación

**7. Herramientas para el análisis de datos**

7.1 Entornos de desarrollo - IDE

7.2 Python

7.3 Librerías

1. **INTRODUCCIÓN**

En este componente se abordarán los conceptos y fundamentos de la inteligencia de negocios, con el fin de realizar el modelado, análisis y preparación de datos. Para una comprensión más detallada, se expone el siguiente video:

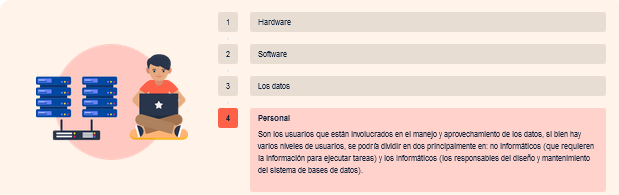
[ir al sitio].

AGREGAR VIDEO

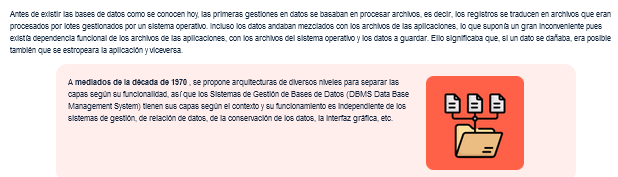
1. **DESARROLLO DE CONTENIDOS:** 
   * + 1. **Tecnologías de información**

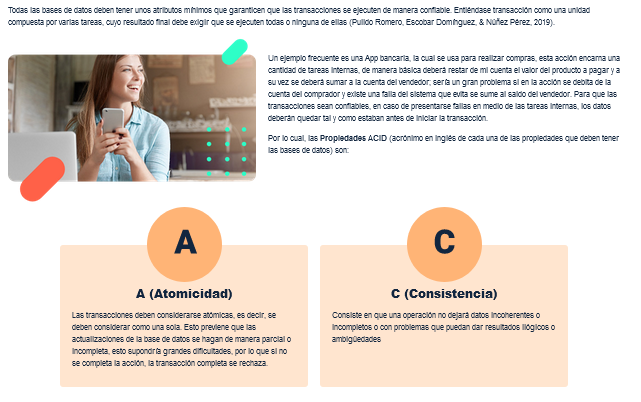
Las bases de datos son esa colección de datos integrados bajo un contexto o dominio, que contienen datos estructurados que reflejan relaciones, restricciones, validaciones y semánticas que reflejan las condiciones de negocio. Esta colección usa la computación para su almacenamiento y procesamiento.

**Los sistemas de base de datos están conformados por los siguientes elementos básicos**:



* 1. **Metodologías de diseño y normalización Principios de ACID**

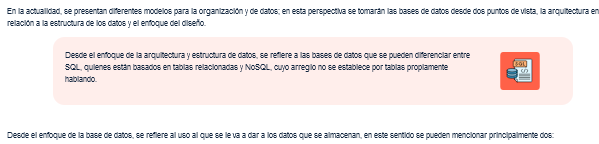




Escala de tiempo

Descripción generada automáticamente

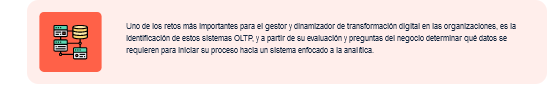
**1.2 Técnicas de almacenamiento de datos y consultas**





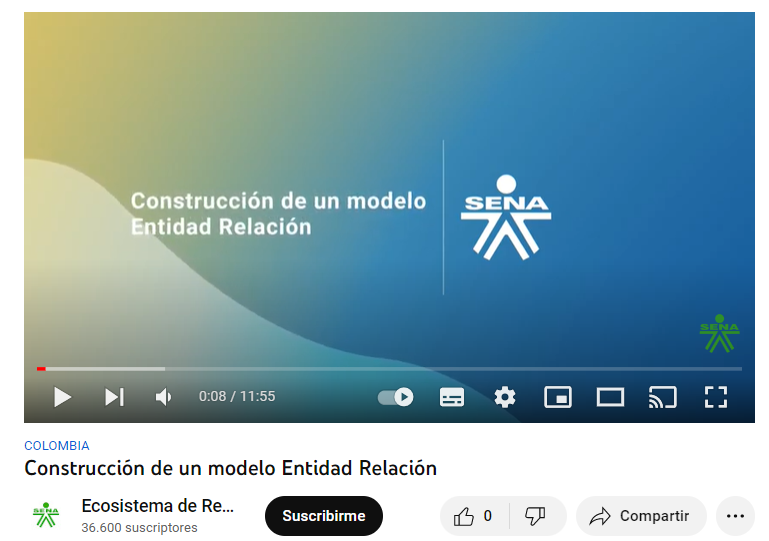
Se componen de diversas tablas que se asocian entre ellas, según las condiciones del negocio, en este sentido se determina qué entidad, mediante la definición de las cosas y actores que intervienen en el sistema, ejemplo: empleado, cliente, sede, producto, categorías, etc. Las entidades son las tablas, y cada entidad tiene atributos tales como nombre, edad, fecha de nacimiento, nombre de producto, presentación, precio, etc. Los atributos serán entonces los campos que tendrán las tablas.

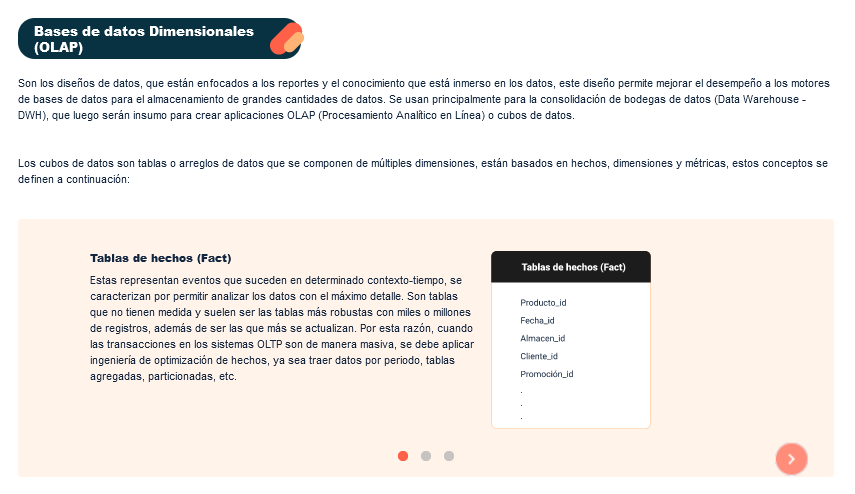
Cada entidad se relaciona con otra tabla a partir de un atributo o campo en común. Este tipo de bases de datos se emplea para sistemas transaccionales o desarrollo de software donde son aplicaciones o desarrollos que dan solución a un proceso específico del negocio, como sistema de facturación, sistema contable, sistemas de inventarios y compras, etc. Se denomina, **OLTP** a lo que en español se nombra como, “Procesamiento de Transacciones En Línea”.



En el siguiente video se presenta un ejemplo de creación de un modelo entidad relación:

(Ir al Sitio)







Texto

Descripción generada automáticamente con confianza media

**1.3 Estándares técnicos**

Tanto para el desarrollo de software como para el diseño de bases de datos, se sugieren unos estándares o convenciones que, si bien no es una norma, se recomienda seguir estas buenas prácticas, en especial porque los proyectos no deben depender de personas, lo que significa que otro profesional que llegue, pueda sentirse familiarizado con los estándares y documentación.

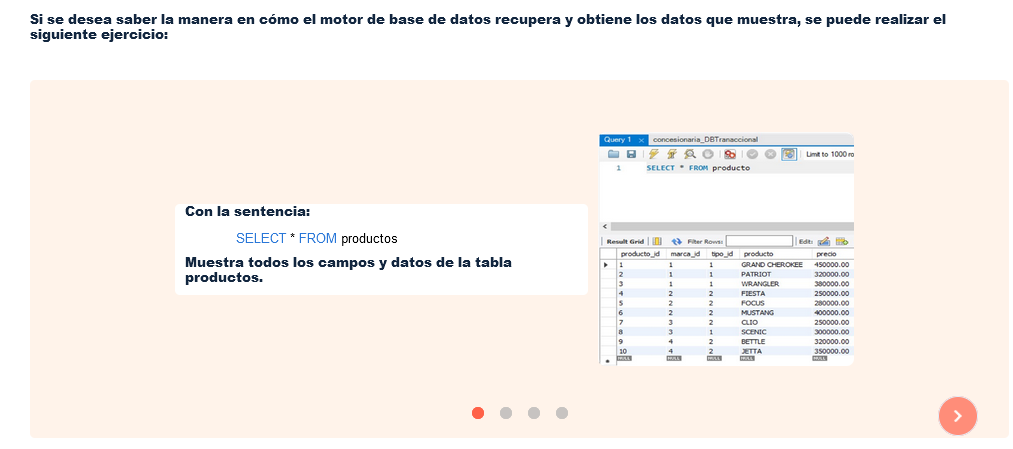
Texto

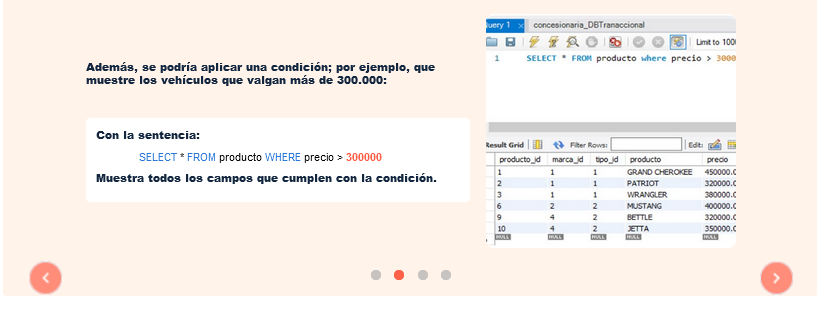
Descripción generada automáticamente

Opcionalmente, los nombres pueden estar en inglés, pues la globalización y los trabajos remotos son una realidad y el inglés es el estándar mundial para la programación y desarrollo de bases de datos y aplicaciones

**1.4 Ordenamiento de datos, indexación y recuperación**

****





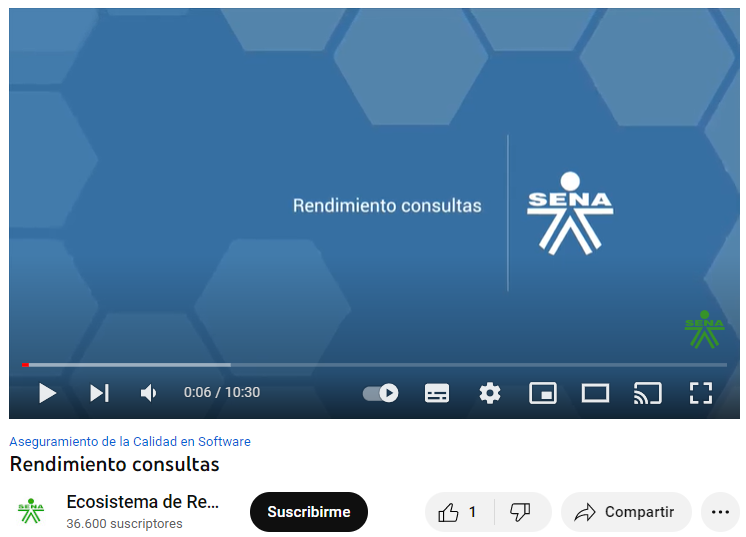
Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Interfaz de usuario gráfica, Texto, Aplicación, Sitio web

Descripción generada automáticamente

Para la manipulación y consulta de datos, es primordial conocer el lenguaje estándar de consulta de las bases de datos relacionales, se invita para que se amplíe el conocimiento sobre este tema en el siguiente video Rendimiento consultas:



**2. Preparación de datos**

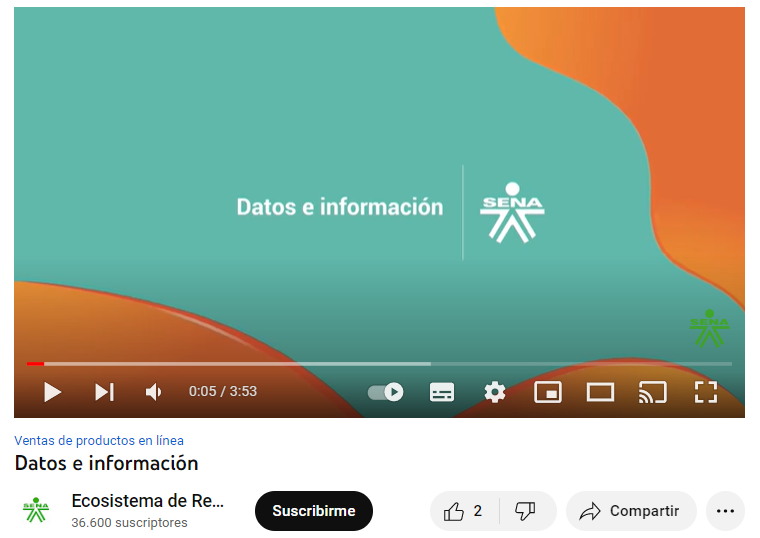


**Interfaz de usuario gráfica

Descripción generada automáticamente**

Bajo el contexto técnico, los datos en la era moderna se guardan a través de máquinas digitales a partir de la implementación de Bases de datos (BD).

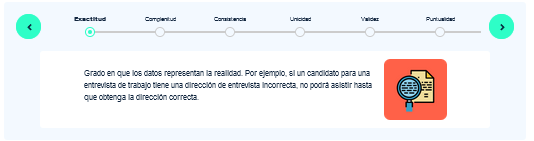
##### Para ampliar información se recomienda ir al siguiente video donde se profundiza en la diferencia entre dato e información:



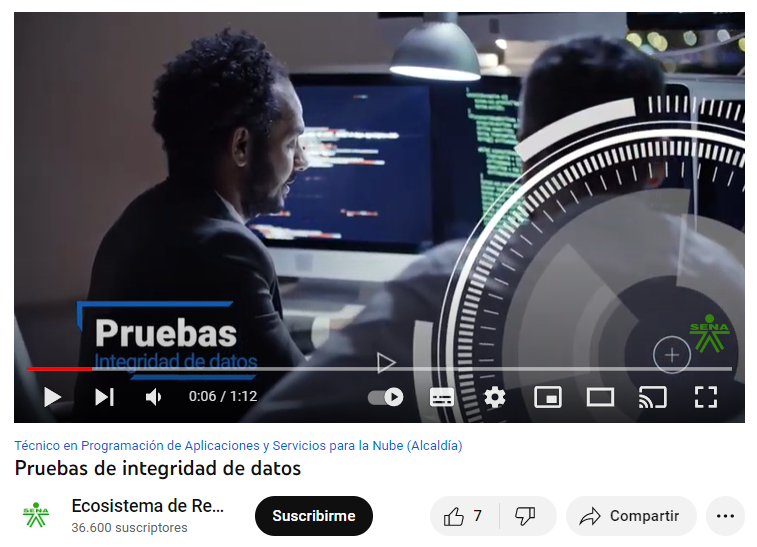
**2.1 Entendimiento de la data**

****

A continuación, se presenta la descripción de cada una de estas dimensiones, que son medibles y que definen la calidad de los datos o Data Quality:



Además, es necesario reconocer la importancia de las pruebas de integridad de datos:



**2.2 Detección de errores y datos faltantes**



Calendario

Descripción generada automáticamente



**2.3 Identificación de variables importantes**



**2.4 Dataset**

**Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente**

Gráfico, Gráfico de embudo

Descripción generada automáticamente

La mayoría de los datos generados por todos los usuarios están clasificados como no estructurados, esto indica que se componen de información difícil de clasificar y ordenar, conformada por archivos sueltos o producidos de manera general como documentos, videos y audios.



Interfaz de usuario gráfica

Descripción generada automáticamente

Interfaz de usuario gráfica

Descripción generada automáticamente

Diagrama

Descripción generada automáticamente

**3.3 Herramientas de administración**

Las herramientas disponibles para la implementación de la inteligencia de negocio en las organizaciones, son amplias y cada vez nuevas marcas y técnicas se disputan el mercado.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Existen otras más, soportadas con grandes marcas como IBM, Google, Teradata, SAP y otras que si bien no son tan mencionadas en el mercado, son igual de poderosas y confiables.

**3.4 Técnicas de solución de problemas (modelación de datos)**

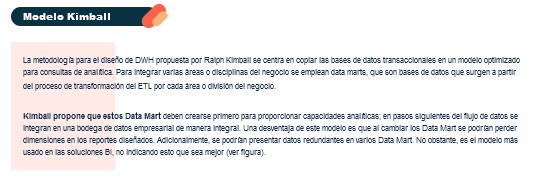
Si bien existen varias metodologías, todas coinciden en una manera genérica para la solución de un proyecto analítico propuesto por Davenport, con algunas variables.

Esta metodología tiene tres grandes etapas:



La modelación de las soluciones BI, depende de las reglas de negocio y los requisitos o problemas a resolver. En los apartados siguientes se detallan las arquitecturas disponibles y el proceso del tratamiento de datos a partir de una mirada de análisis e inteligencia de negocio.

**3.5 Metodologías de análisis (Kimball, Inmon)**

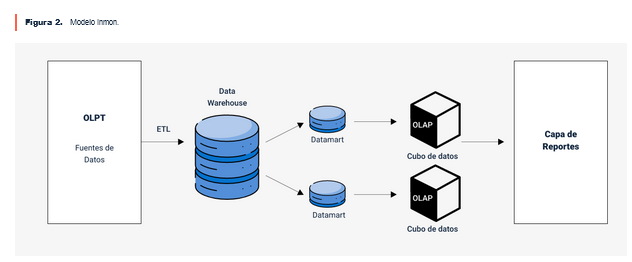
En lo relacionado a la planeación y diseño de las bodegas de datos (DWH), se plantean dos arquitecturas:****

**Diagrama

Descripción generada automáticamente**

**Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente**



**3.6 Verificación de valores y escalas**

**

**

*Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente*

En esta tabla se presentan ejemplos de disparadores, para una mejor comprensión.

**4. Análisis exploratorio de datos**

**Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente**

**4.1 Estadística descriptiva y estadística inferencial**

Son los métodos mediante los cuales se presenta la información y pueden clasificarse en:

Texto

Descripción generada automáticamente

## 4.2 Población y muestra

Para la aplicación de estadística inferencial, se tratarán algunos parámetros matemáticos para la aplicación de este modelo estadístico que se ilustran a continuación:

Imagen que contiene texto

Descripción generada automáticamente

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.

Texto

Descripción generada automáticamente

**4.3 Escalas de medida y clasificación de variable**

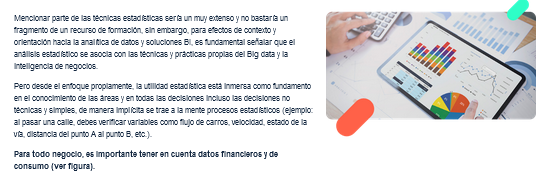
Las escalas de medición son procesos de comparación y dimensionalidad, que da cuenta de un valor que signifique sus proporciones. Una variable, es un elemento que cambia y que al hacer parte de un conjunto puede afectar.

Las escalas se convierten en algo real a través de las preguntas que se utilizan para recolectar la información aplicando la escala. (Domínguez, 2017). Existen cuatro escalas primarias de medición: nominal, ordinal, intervalo y radio; algunas escalas más sofisticadas como las escalas multi contenido o multipropósito, en la siguiente tabla se describen las categorías.

Tabla

Descripción generada automáticamente

**4.4 Técnicas de análisis estadístico**

****

**Gráfico, Gráfico de barras, Gráfico en cascada

Descripción generada automáticamente**

**Texto, Carta, Correo electrónico

Descripción generada automáticamente**

****

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.

**5. Métodos para hacer análisis exploratorio de datos**

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Imagen que contiene Aplicación

Descripción generada automáticamente

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Texto, Correo electrónico

Descripción generada automáticamente

**5.4 Reglas de negocio**

Desde el punto de vista de programación, el cual debe coincidir con el negocio, son las condiciones asociadas a las tareas que forman los procesos. Las reglas no son la normatividad de las empresas, se refiere más a los requisitos o necesidades y cómo los sistemas de información satisfacen estos requisitos bajo las reglas y condiciones técnicas necesarias.

**Las reglas de negocio se basan en las políticas de las organizaciones, en metodologías ágiles como el Scrum, las reglas de negocio se plasman en historias de usuarios.**

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

Imagen que contiene Diagrama

Descripción generada automáticamente

A estas reglas de negocio, a menudo se les denomina también **Lógica de negocio**.



**Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente**

**5.7 Programación de estructuras no lineales, desnormalización, series y dataframes**

En una estructura lineal, cada elemento sólo puede ir enlazado al siguiente o al anterior. A las estructuras de datos no lineales se les llama también estructuras de datos multienlazadas y tiene las siguientes características

Diagrama

Descripción generada automáticamente

**Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente**

Por eso es importante tener presente las siguientes recomendaciones entregadas en el video Manipular los datos:



****

Existen varios motores de bases de datos con núcleo SQL, entre los más conocidos MySQL de uso libre y otra distribución licenciada por Oracle.

En el siguiente video se presentan las ventajas de Mysql:



SQL server pertenece a la casa de Microsoft, con sus herramientas integradas que dan mucho poder para el almacenamiento y proceso de datos.

En el siguiente video se presentan los comandos de SQL para la creación de una base de datos y para la realización de consultas en esta:



**5.10 No SQL**

Son un tipo de bases de datos cuyo almacenamiento de datos no se realiza en tablas propiamente dicha. Presentan una estructura particular tipo JSON o BSON, que consiste en un arreglo de datos por jerarquías, tiene la ventaja sobre las bases de datos relacionales que su desempeño de búsquedas y cargue de datos son más veloces, además, se podría definir que son más flexibles es sus datos y cambios de estructuras que se puedan presentar. Para el manejo de Big Data y grandes cantidades de datos son muy empleadas.

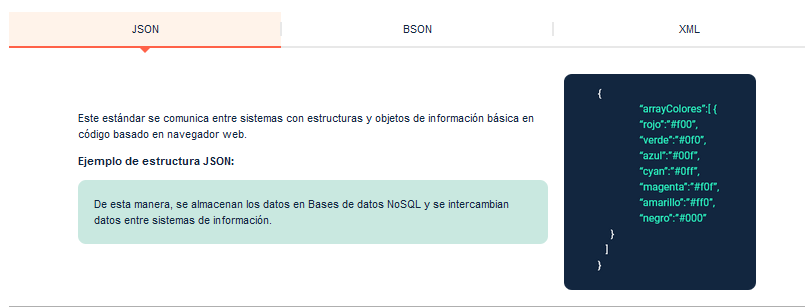
En siguiente video se explican las ventajas de las bases de datos no relacionales o No SQL:



**5.11 JSON, BSON y XML**

Uno de los desafíos técnicos es la integración entre sistemas de información y la manera en cómo enviar y recibir datos de otras aplicaciones.

Para ello se desarrollaron diversas maneras, entre las más comunes son:

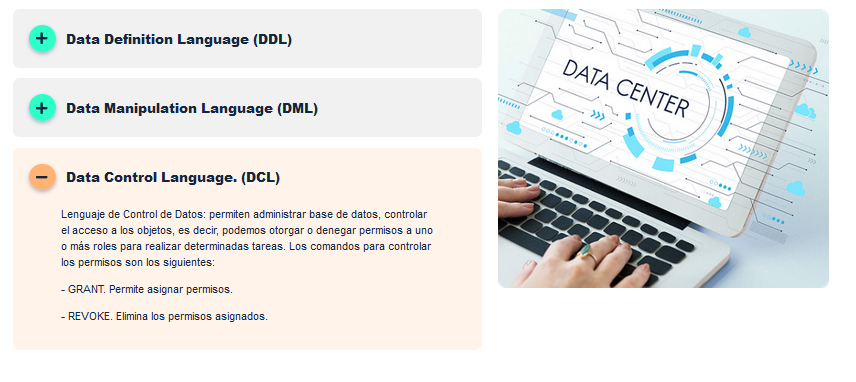


Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

**5.12 DDL, DML, DCL**

Las bases de datos emplean diferentes tipos de lenguajes y también de allí se derivan archivos necesarios para que los motores de bases de datos funcionen y conserven la información física y lógicamente. Estas son:



**6. Estructuras y componentes de analítica de datos**

Icono

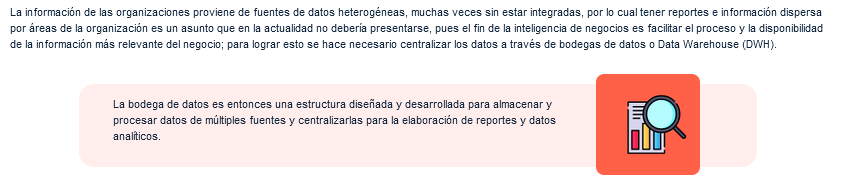
Descripción generada automáticamente con confianza bajaA continuación, este apartado se emplea para dar una breve introducción a elementos muy importantes a hora de aplicar inteligencia de negocios. Se describe a modo conceptual, algunos elementos de las bodegas de datos y sus topologías.

La estructura completa de la analítica tiene dos extremos donde fluyen y se transforman los datos, desde los orígenes, pasando por un proceso de carga al sistema analítico, limpieza de datos, optimización de datos para ser guardados de forma consolidada y completa en un arreglo de bases de datos denominado Bodega de datos, pero acá no acaba el flujo, el final es la explotación de datos traducidos en reportes, cubos de datos u otros usos como minería de datos.

Con el siguiente video, se amplia información sobre analítica de datos, con lo cual puede despejar algunas dudas:



**6.1. Bodega de datos**



Texto

Descripción generada automáticamente

**6.2 Tipos estrella**

Es aquella que cuyas dimensiones se relacionan directamente con la tabla de hechos, se representa de la siguiente manera

Diagrama

Descripción generada automáticamente

Vale la pena señalar que muchas tablas giran en torno a la tabla principal, conocida como la tabla de hechos, la cual generalmente contiene las claves primarias de las dimensiones asociadas, así como las cifras o medidas obtenidas previamente en el proceso ETL.

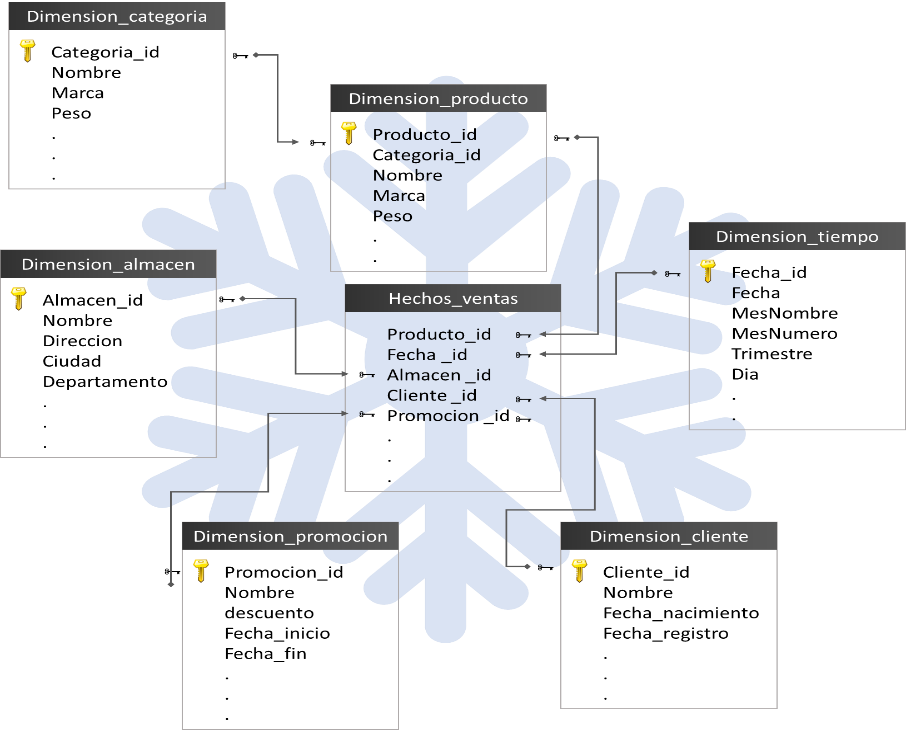
**6.3 Copo de nieve**

La topología copo de nieve o *Snowflake* se presenta frecuentemente, cuando las dimensiones cuentan con otras familias o categorías que obligan a incluir subdimensiones para completar la información.

Ejemplo de esto es cuando existe una dimensión **Productos**, es muy frecuente que exista la dimensión **Categorías**. Esta última se conecta a la tabla de hechos a través de **Productos** que está en medio. Se representa de la siguiente manera (ver figura):

**Figura 7**

*DWH Snowflake*



En este caso, al igual que con la topología de estrella, los hechos se relacionan con las dimensiones a partir de las claves de las tablas que lo circundan, sin embargo, en este caso se pueden presentar más niveles donde una dimensión pueda contener una clave foránea de otra tabla. En ese sentido, como se ve en el ejemplo, la dimensión\_categoria se asocia a los hechos a través de una tabla en medio llamada dimension\_producto.

**6.4 Constelación**

Similar a la topología copo de nieve, sin embargo, se presenta cuando hay más de dos niveles de relación. Es decir, hay más de una tabla en medio entre una dimensión y la tabla de hechos. Esta arquitectura no es muy común y es poco eficiente, pues ya los diseños dimensionales se van desdibujando un poco y comienzan más a parecerse a estructuras transaccionales que multidimensionales.

Al sumar muchos datos en tablas de hechos y dimensiones, más el procesamiento que implica la lectura y escritura de datos con muchas relaciones de tablas y todo de manera masiva, no hace muy veloz estas arquitecturas, ocasionalmente algunas soluciones traen las tablas del sistema transaccional a las bodegas de datos sin procesos ETL, implicando mucha carga para los sistemas de reportes ocasionando su colapso en algunas herramientas de visualización de datos.

**7. Herramientas para el análisis de datos**

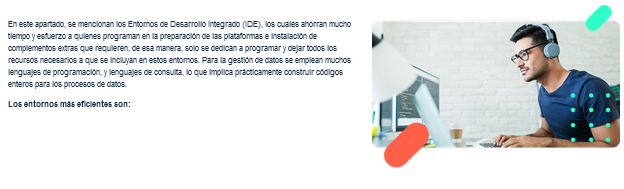
Texto

Descripción generada automáticamente con confianza baja

Tabla

Descripción generada automáticamente

## 7.1 Entornos de desarrollo - IDE

****

**Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente**

**7.2 Python**

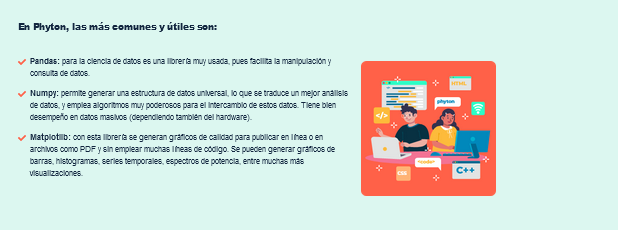
Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamenteActualmente, es el lenguaje de programación que lidera los desarrollos basados en gestión de datos. En relación con otros lenguajes presenta curva de aprendizaje rápido, cuenta con múltiples librerías que expanden su capacidad y muchos sistemas de manejo de datos e inteligencia artificial lo emplean para codificar sus funcionalidades.

Para desarrollar aplicaciones y realizar minería de datos basado en inteligencia artificial puede emplear alguno de los IDE anteriormente conocido u otros como *Code Skulpor* <https://py3.codeskulptor.org/>

**7.3 Librerías**

Las librerías en los lenguajes de programación son un conjunto de archivos que contienen códigos de programas o funcionalidades que hacen tareas concretas y repetitivas para facilitar la programación. Las librerías se categorizan por el tipo de funcionalidades que presentan, algunas son de tipo estadístico, otras para la visualización de datos, otras para operaciones matemáticas especializadas, etc.



Síntesis

Interfaz de usuario gráfica

Descripción generada automáticamente

Diagrama

Descripción generada automáticamente

1. **ACTIVIDADES DIDÁCTICAS**

|  |  |
| --- | --- |
| DESCRIPCIÓN DE ACTIVIDAD DIDÁCTICA | |
| Nombre de la Actividad | Modelamiento, análisis y preparación de datos. |
| Objetivo de la actividad | Identificar los conceptos sobre modelamiento, análisis y preparación de datos con el fin de reforzar los conocimientos adquiridos en el componente. |
| Tipo de actividad sugerida | https://lh6.googleusercontent.com/FvABbrwgrnxeHtlHBpt0FdRjJH5VpWBdLbK0Y-NyHF0MxmgecZxbT_QLjV8v2h_0Gr5ETSV7C0ihbVBtf_VBqiBSiEmxWXfjkqpV-5uTXN3HQPFv2D2_nnl_RHrCosO6f6A5mfAWEenp2HD8W1mNoBi4amrW5VkKn7N1ZGfcpGbi1gj0WLQjzi5eb6uE_Sz7u4uqGA |
| Archivo de la actividad  (Anexo donde se describe la actividad propuesta) | Anexo\_CF13\_ActividadDidactica |

**MATERIAL COMPLEMENTARIO:**

|  |  |  |  |
| --- | --- | --- | --- |
| Tema | Referencia APA del Material | Tipo de material  (Video, capítulo de libro, artículo, otro) | Enlace del Recurso o  Archivo del documento o material |
| 1. Modelamiento y gestión de bases de datos | Quintana, G. (2014). Aprende SQL. Universitat Jaume I. Servei de Comunicació i Publicacions. | Libro | <https://elibro-net.bdigital.sena.edu.co/es/lc/senavirtual/titulos/53252> |
| 1.2 Técnicas de almacenamiento de datos y consultas | Ecosistema de Recursos Educativos Digitales SENA. (2021). Aplicando el MER con herramienta Día || Cardinalidad modelo entidad relación. | Video | <https://youtu.be/KcORNp2A3yg> |
| 2.2 Detección de errores y datos faltantes | Codificandobits. (s.f). Mapa paso a paso manejo datos faltantes. Blog. | Guía | <https://www.codificandobits.com/descargas/dl_202100618_mapa_pasos_manejo_datos_faltantes.pdf> |
| 2.3 Identificación de variables importantes | Sotaquirá, M. (2021). Guía completa para el manejo de datos faltantes. | Artículo | <https://www.codificandobits.com/blog/manejo-datos-faltantes/> |
| 3.7 Procedimientos almacenados y funciones | Calbimonte, D. (2019). Funciones frente a los procedimientos almacenados en SQL Server. | Tutorial | <https://www.sqlshack.com/es/funciones-frente-a-los-procedimientos-almacenados-en-sql-server/> |
| 4. Análisis exploratorio de datos | Codificandobits. (s.f). Guía paso a paso análisis exploratorio. | Mapa conceptual | <https://www.codificandobits.com/descargas/dl_202100611_mapa_pasos_analisis_exploratorio.pdf> |
| 5.8 Algebra relacional | Cidecam. (2021). Algebra Relacional. | Artículo | <http://cidecame.uaeh.edu.mx/lcc/mapa/PROYECTO/libro14/33_algebra_relacional.html> |
| 7.3 Librerías | Ecosistema de Recursos Educativos Digitales SENA. (2021). Manejo de datos con Pandas. | Vídeo | <https://youtu.be/J9_xVhV_pfs> |

1. **GLOSARIO:**

|  |  |
| --- | --- |
| TÉRMINO | SIGNIFICADO |
| Información Descriptiva | Se refiere a las historias que los datos cuentan, se refiere a un histórico de eventos y resultados. |
| Información predictiva | Se refiere a los hechos probables que sucederán, esto se realiza basado en datos descriptivos, es decir, datos históricos y procesos matemáticos y/o modelación de Machine Learning. |
| Joins | Proceso de selección de datos de varias tablas y unirlos en una vista o nueva tabla. Una instrucción de “SQL JOIN” en un comando Select que combina columnas entre una o más tablas en una base de datos. |
| Lenguaje DAX | Lenguaje específico para gestión de datos creado por Microsoft (Data Analysis Expressions). Se emplea en colecciones de datos en aplicaciones como Excel, Analysis Services y Power BI. |
| Lenguaje R | Lenguaje de programación para la gestión de datos. Es un lenguaje interpretado que ejecuta las instrucciones directamente sin previa compilación. |
| Normalización | La estructura organizada en datos relacionales que cumplen unas reglas de normalización que garantizan la integridad, calidad y optimización en la base de datos. |
| Procesadores | CPU (Unidad central de proceso), es el componente del computador y otros dispositivos programables, que interpreta las instrucciones contenidas en los programas y procesa los datos. |
| Procesamiento por lotes | Al tener muchas cargas de datos y procesamientos, es necesario reunir recursos para que se ejecuten de manera independiente optimizando recursos, de esta manera las tareas se completan periódicamente de manera repetitiva |
| RAM | Es la memoria temporal o de corto plazo de las computadoras, es la memoria principal de trabajo, los programas y datos se cargan allí para que trabajen más rápidamente. |
| Script | se refiere a fragmentos de código de programación que pueden ejecutar una o varias funciones. |
| Sistema operativo | Es el software principal de las computadoras, se emplea como plataforma para gestionar las aplicaciones, recursos del hardware y entornos gráficos y funcionales. |
| Tabular | En estadística, son la recopilación y procesamiento de la información capturada de los instrumentos disponibles al momento de realizar encuestas, toma de datos y otras. |
| TI | (IT) Abreviatura de Tecnología de la información. |

1. **REFERENCIAS BIBLIOGRÁFICAS:**

Banco de la República. (2022). *Sistema de información económica de la Gerencia Técnica*. <https://totoro.banrep.gov.co/analytics/saw.dll?Portal&PortalPath=%2Fshared%2FDashboards_T%2FD_Estad%C3%ADsticas%2FEstad%C3%ADsticas&NQUser=publico&NQPassword=publico123&lang=es&page=Precios%20e%20inflaci%C3%B3n>

Conza, A. (2018). *Estándares para diseño e implementación de bases de datos relacionales*. http://adrianconza.com/: <http://adrianconza.com/estandares-de-bases-de-datos/>

Curto Díaz, J. (2016). *Introducción al Business Intelligence*. Editorial UOC. <https://elibro-net.bdigital.sena.edu.co/es/lc/senavirtual/titulos/101030>

Díaz, J. C. (2016). *Organizaciones orientadas al dato: transformando las organizaciones hacia una cultura analítica*. Editorial UOC. <https://elibro-net.bdigital.sena.edu.co/es/lc/senavirtual/titulos/58609>

Fernández, J. (2021). Escalas de medición de las variables: nominal, ordinal, intervalo y razón. <https://youtu.be/XNuIqSfCskQ>

Gawande, S. (2020). iCEDQ Torana INC. *6 Dimensions of Data Quality, Examples, and Measurement*. <https://icedq.com/6-data-quality-dimensions>

Ommi, A. K. (2018). *Introduction to Data and Information*. MyCloudWiki: <https://www.mycloudwiki.com/san/data-and-information-basics/>

Pang, A., Markovski, M., & Ristik, M. (2022). *Los 10 principales proveedores de software de análisis y BI, tamaño del mercado y pronóstico del mercado 2021-2026*. Apps Run the World. <https://www.appsruntheworld.com/top-10-analytics-and-bi-software-vendors-and-market-forecast/>

Pulido Romero, E., Escobar Dominguez, O., & Núñez Pérez, J. (2019). *Bases de datos*. Grupo Editorial Patria. Obtenido de <https://elibro-net.bdigital.sena.edu.co/es/lc/senavirtual/titulos/121283>

Velthuis, M. P. (2019). *Calidad de datos*. Ediciones de la U. <https://www-ebooks7-24-com.bdigital.sena.edu.co/?il=9094>

1. **CONTROL DEL DOCUMENTO**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Nombre | Cargo | Dependencia | Fecha |
| Autor (es) | Jaime Hernán Tejada | Experto Temático | Centro de la Industria, la empresa y los servicios - CIES | Noviembre de 2022 |
| Giovanna Andrea Escobar Ospina | Diseñador Instruccional | Sena - Norte de Santander Centro CIES | Noviembre de 2022 |
| Silvia Milena Sequeda Cárdenas | Asesora metodológica y pedagógica | Regional Distrito Capital -Centro de Diseño y Metrología. | Noviembre de 2022 |
|  | Jhon Jairo Rodríguez Pérez | Corrector de Estilo | Regional Distrito Capital - Centro de Diseño y Metrología | Noviembre de 2022 |

1. **CONTROL DE CAMBIOS**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Nombre | Cargo | Dependencia | Fecha | Razón del Cambio |
| Autor (es) | Sandra Paola Morales Paez | Evaluador Instruccional | Regional  Santander  Centro Agroturístico | 3 de junio de 2024 | Adecuaciones a 2024 |
| Claudia Johanna Gómez Pérez | Líder Línea Santander |