



Unidad 2

“Preparación de los datos”

**Documento para evaluar el Saber de la materia de
Administración de base de datos**

**Ingeniería
en
Tecnologías de la Información y Comunicación
Desarrollo y Gestión de software**

**Maestra:
M.C Carlos Alfredo Montes Carnero**

**Elaborado por:
Jonathan Jared Salazar Viera**

Índice

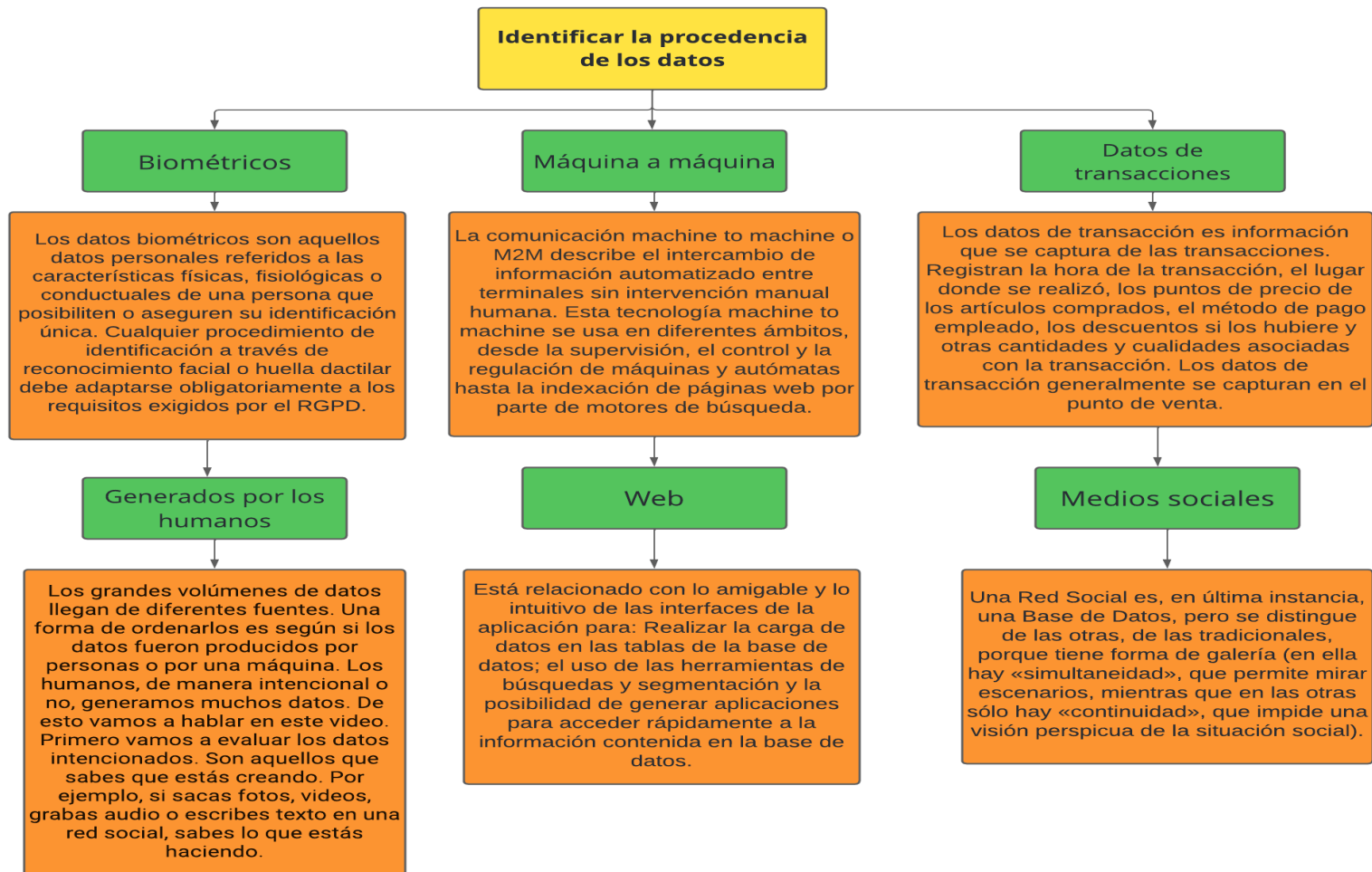
Introducción	3
Desarrollo.....	4
Identificar la procedencia de los datos.....	4
Identificar los tipos de datos en análisis de datos.....	5
Identificar los tipos de fuentes de datos.....	6
Identificar las características de los data warehouse.	7
Identificar los esquemas de data warehouse.....	8
Definir el concepto de limpieza de datos e identificar las técnicas de limpieza de datos. 9	
Definir el concepto de datos objetivo, datos de entrenamiento y datos de prueba.	10
Identificar repositorios de datos y herramientas de generación de datos.	11
Identificar las técnicas de extracción de datos.	12
Identificar las técnicas de transformación de datos.	13
Identificar las técnicas de carga de datos.	14
Conclusiones.....	15
Bibliografía.....	16

Introducción

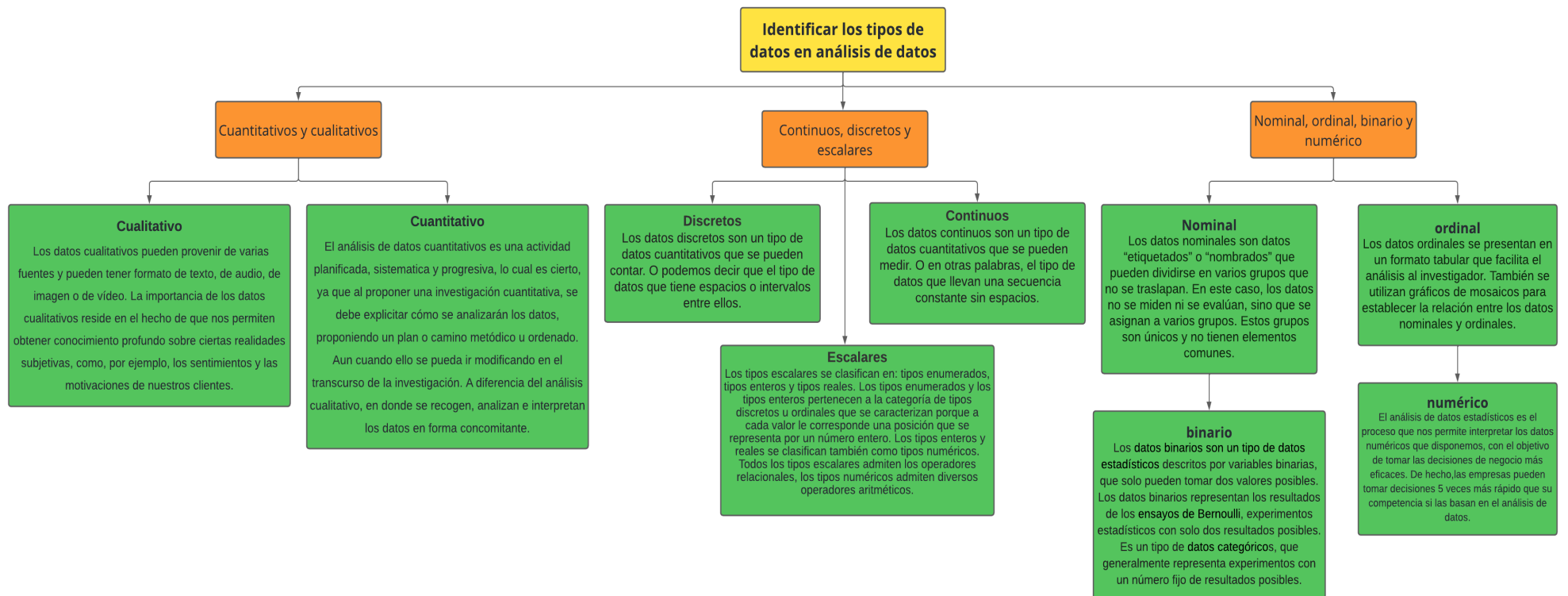
En el siguiente documento se estableció el tema de procedimiento de datos que en los cuales se desprenden varios subtemas que en estos subtemas se tienen estimulados con información en mapa cognitivo de cajas por lo cual A medida que la potencia de los sistemas informáticos se incrementa, la necesidad de información crece proporcionalmente, llevando a un crecimiento cada vez mayor de la recopilación de datos: más casos, más variables y más errores en la entrada de datos. Estos errores son la pesadilla de las predicciones del modelo predictivo, que son el objetivo final del almacenamiento de datos, por lo que existe la necesidad de mantener los datos "limpios". Sin embargo, la cantidad de datos almacenados ha superado de tal forma a la capacidad de comprobar los casos manualmente que resulta vital implementar procesos automatizados para validar los datos.

Desarrollo

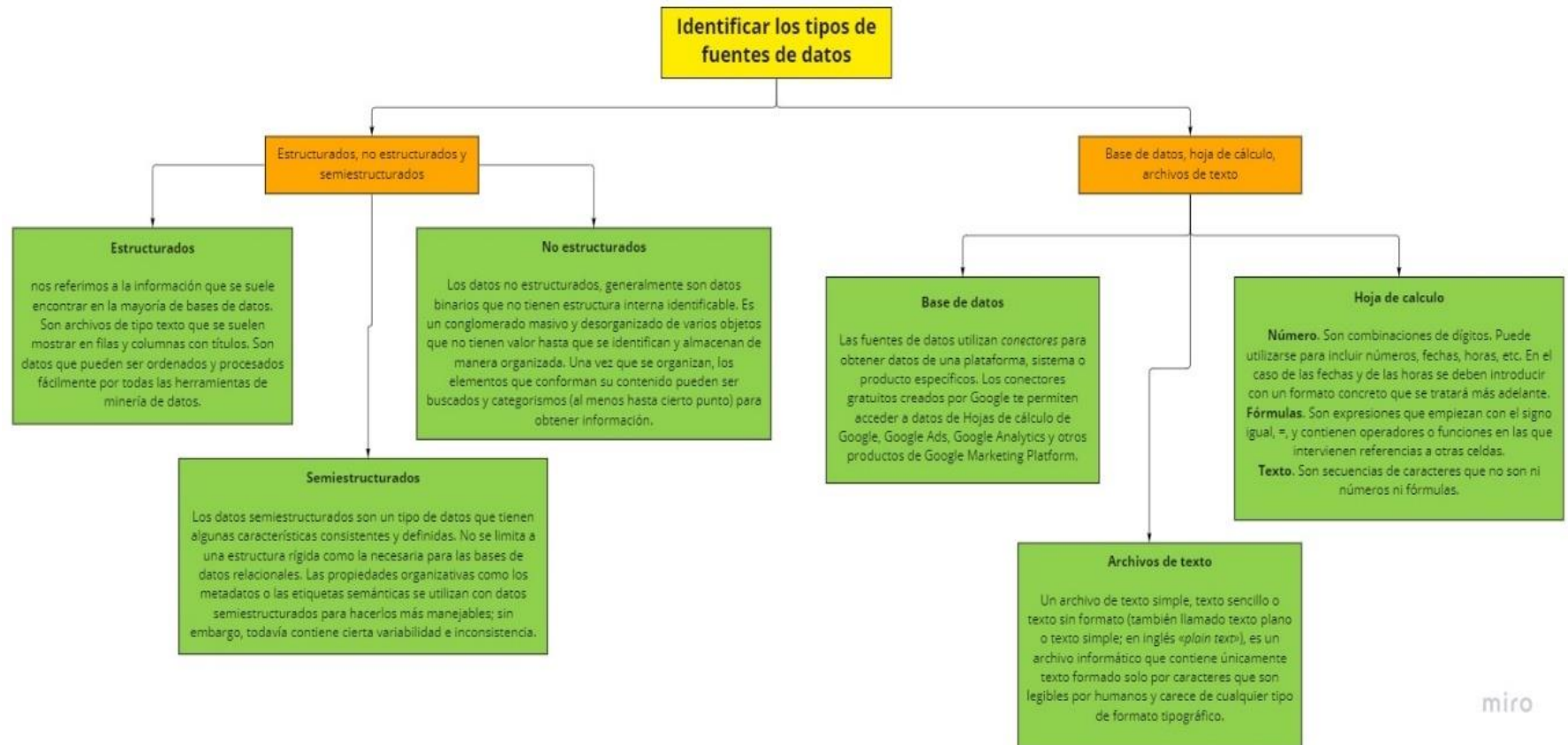
Identificar la procedencia de los datos



Identificar los tipos de datos en análisis de datos

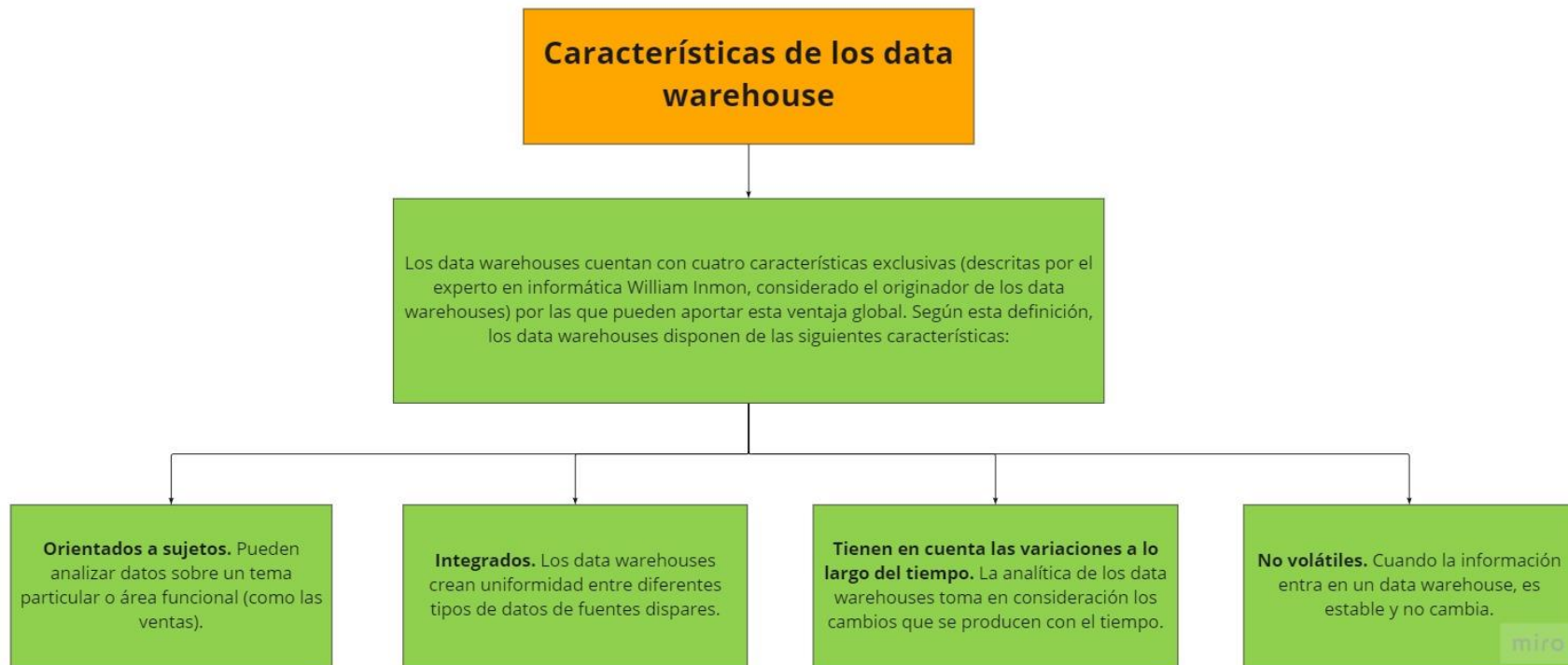


Identificar los tipos de fuentes de datos

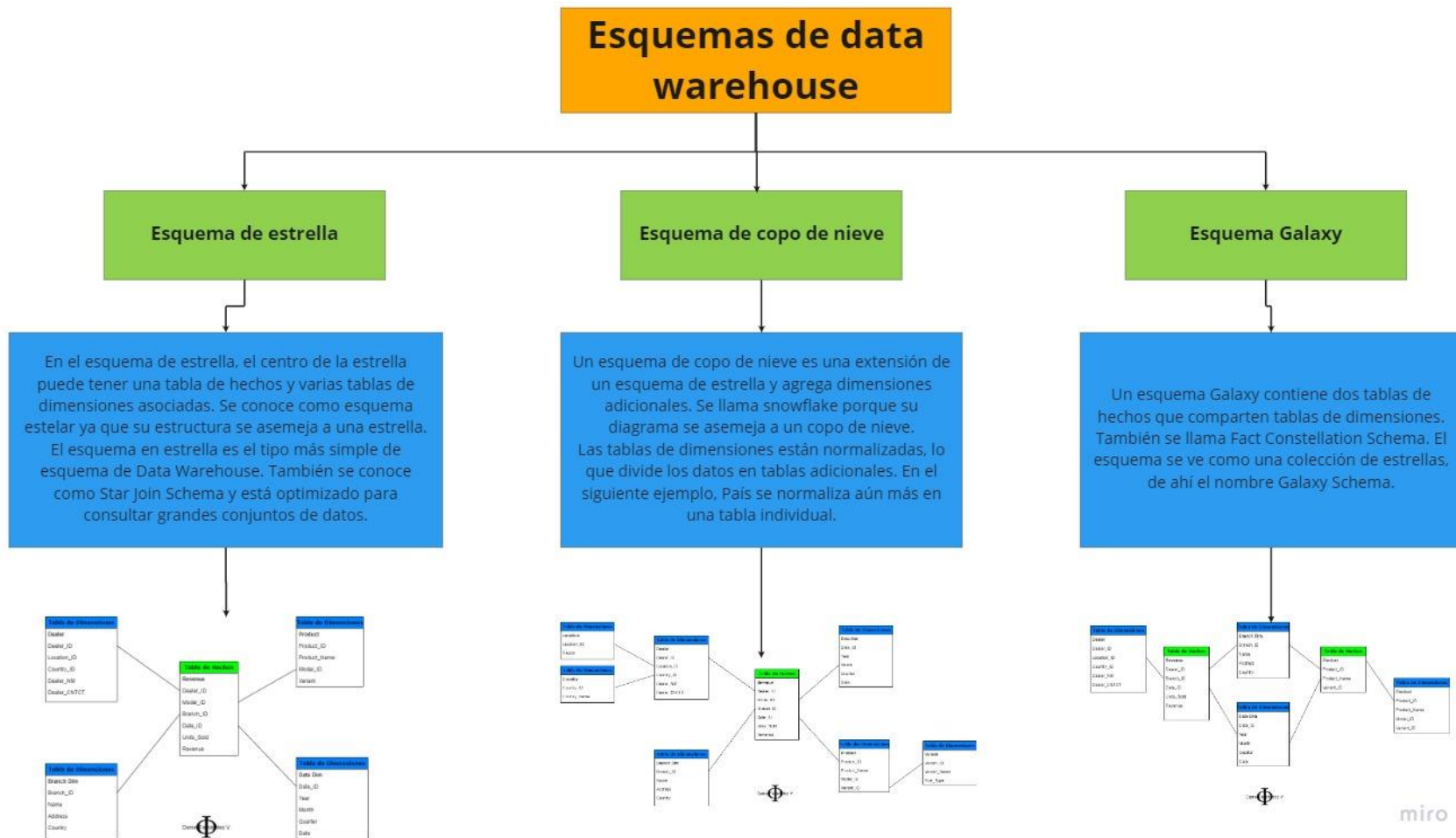


miro

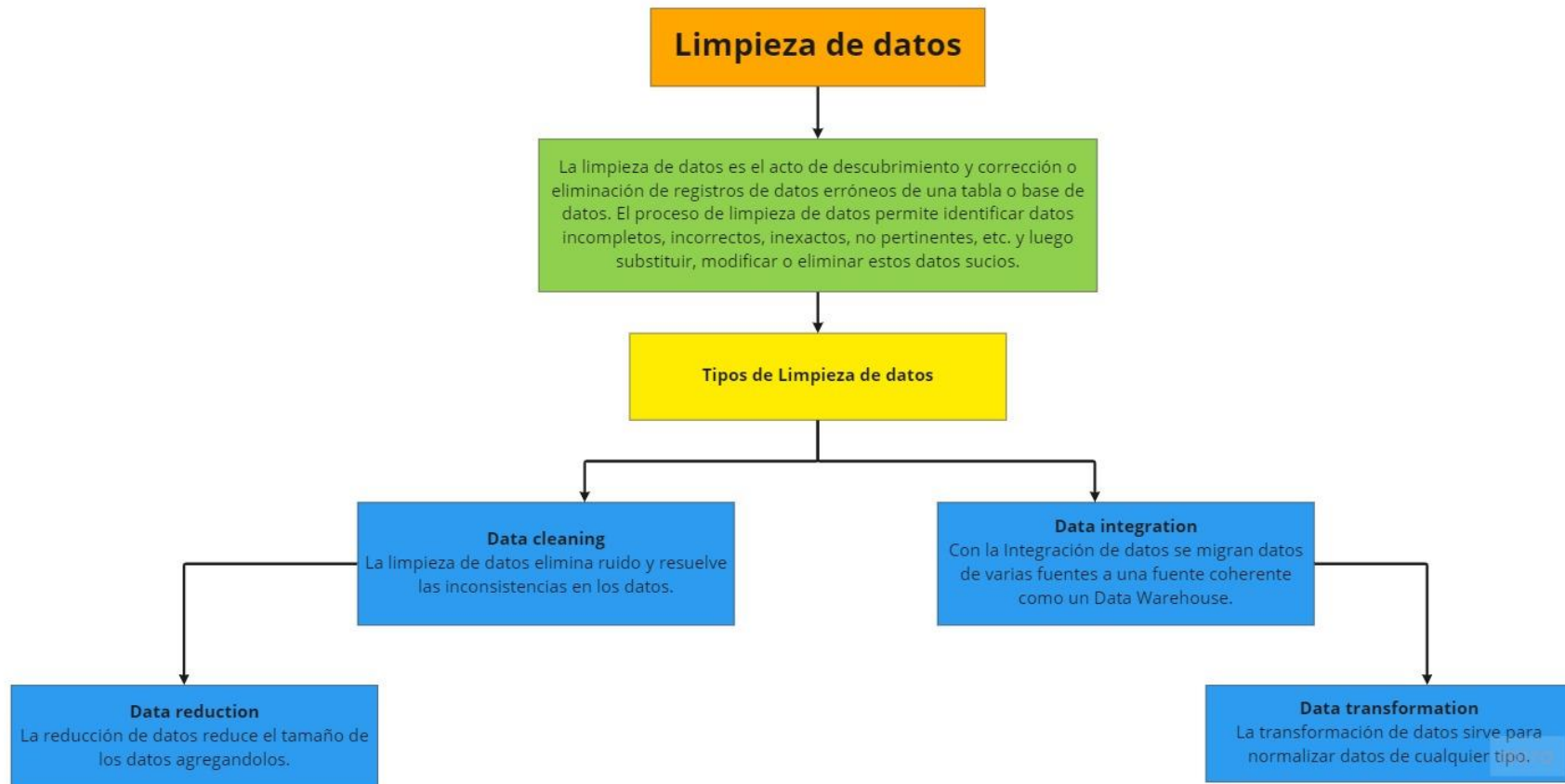
Identificar las características de los data warehouse.



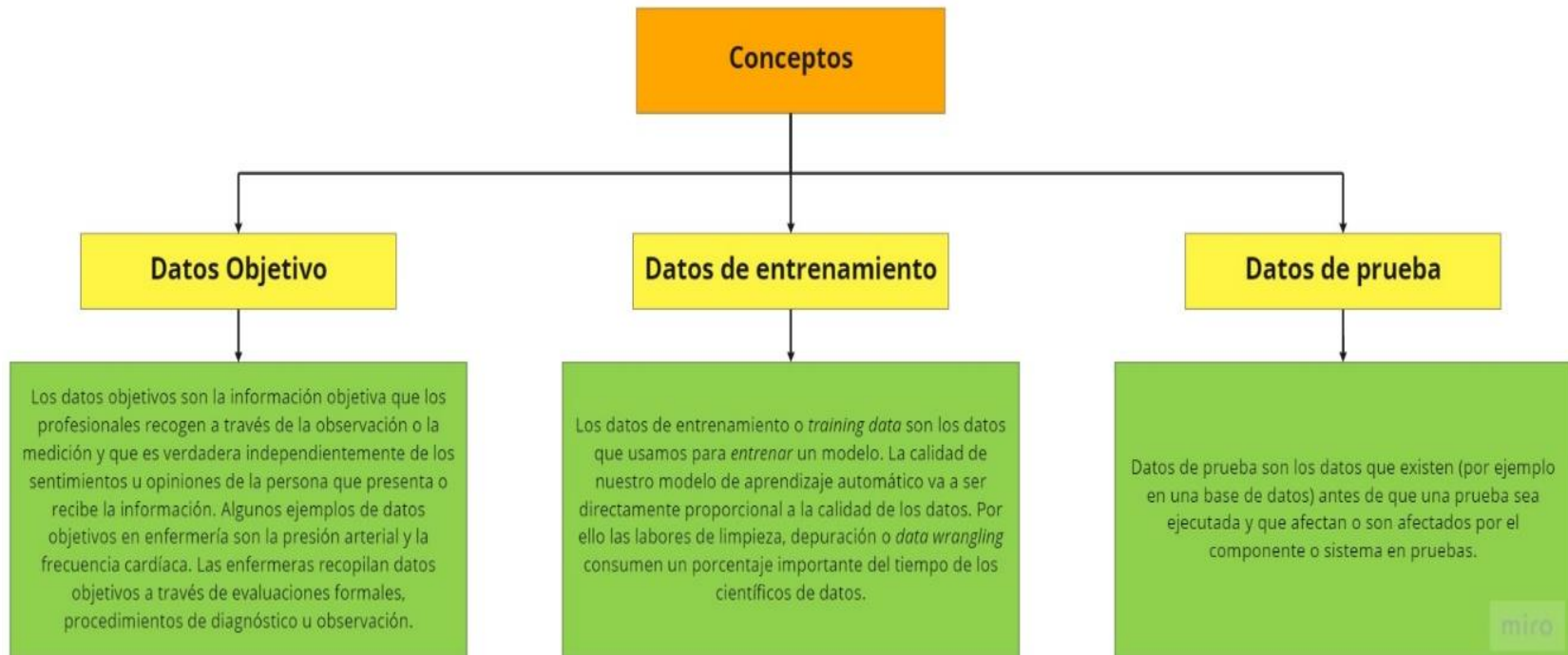
Identificar los esquemas de data warehouse.



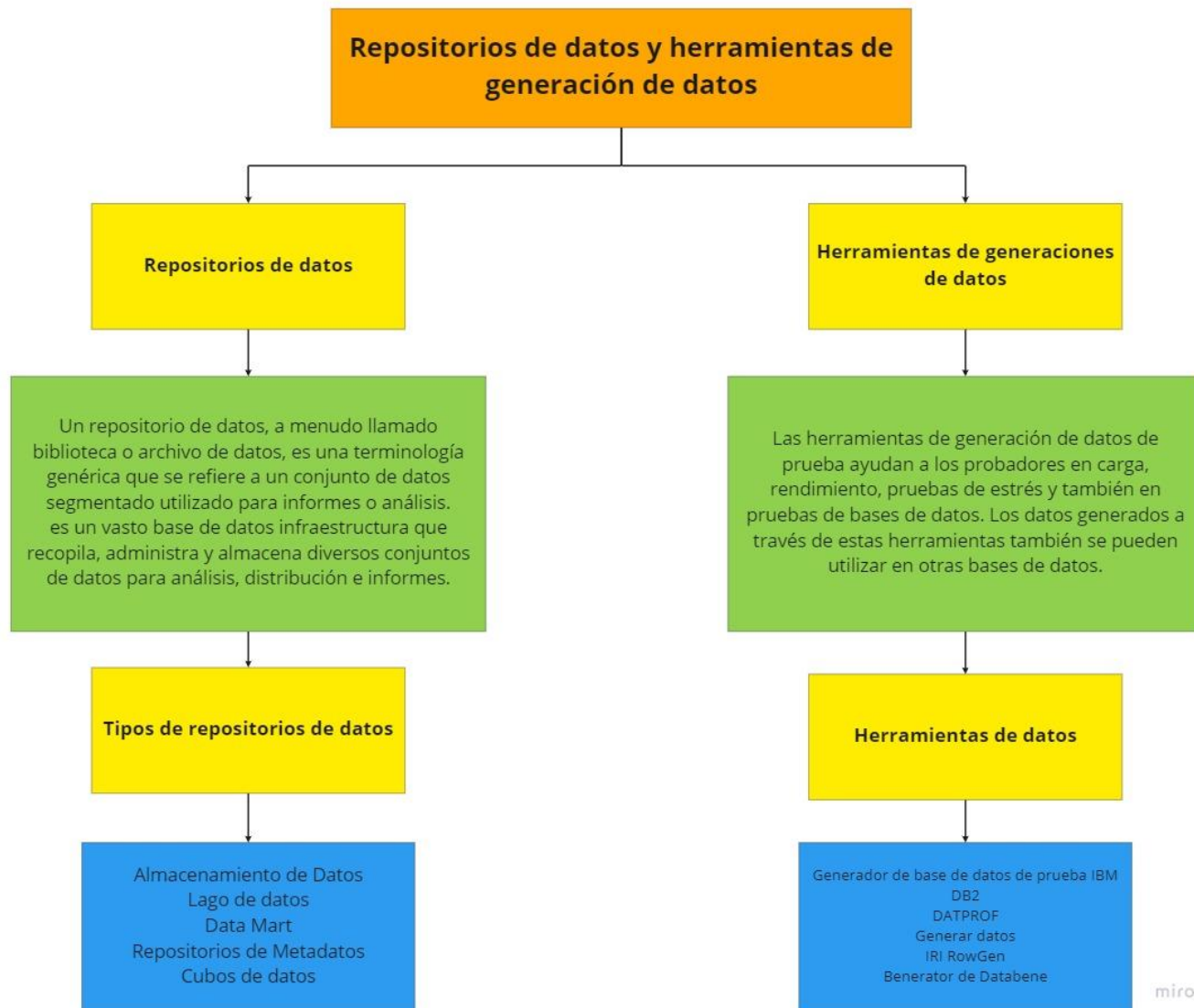
Definir el concepto de limpieza de datos e identificar las técnicas de limpieza de datos.



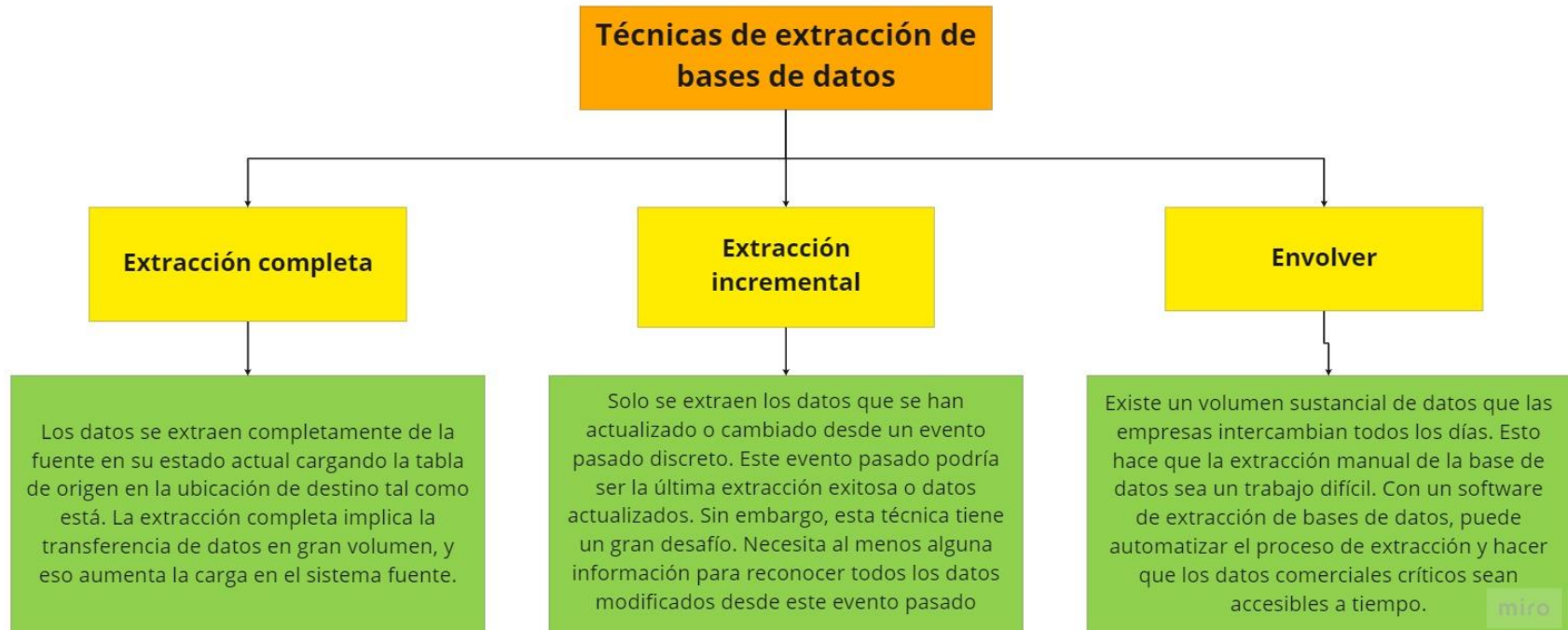
Definir el concepto de datos objetivo, datos de entrenamiento y datos de prueba.



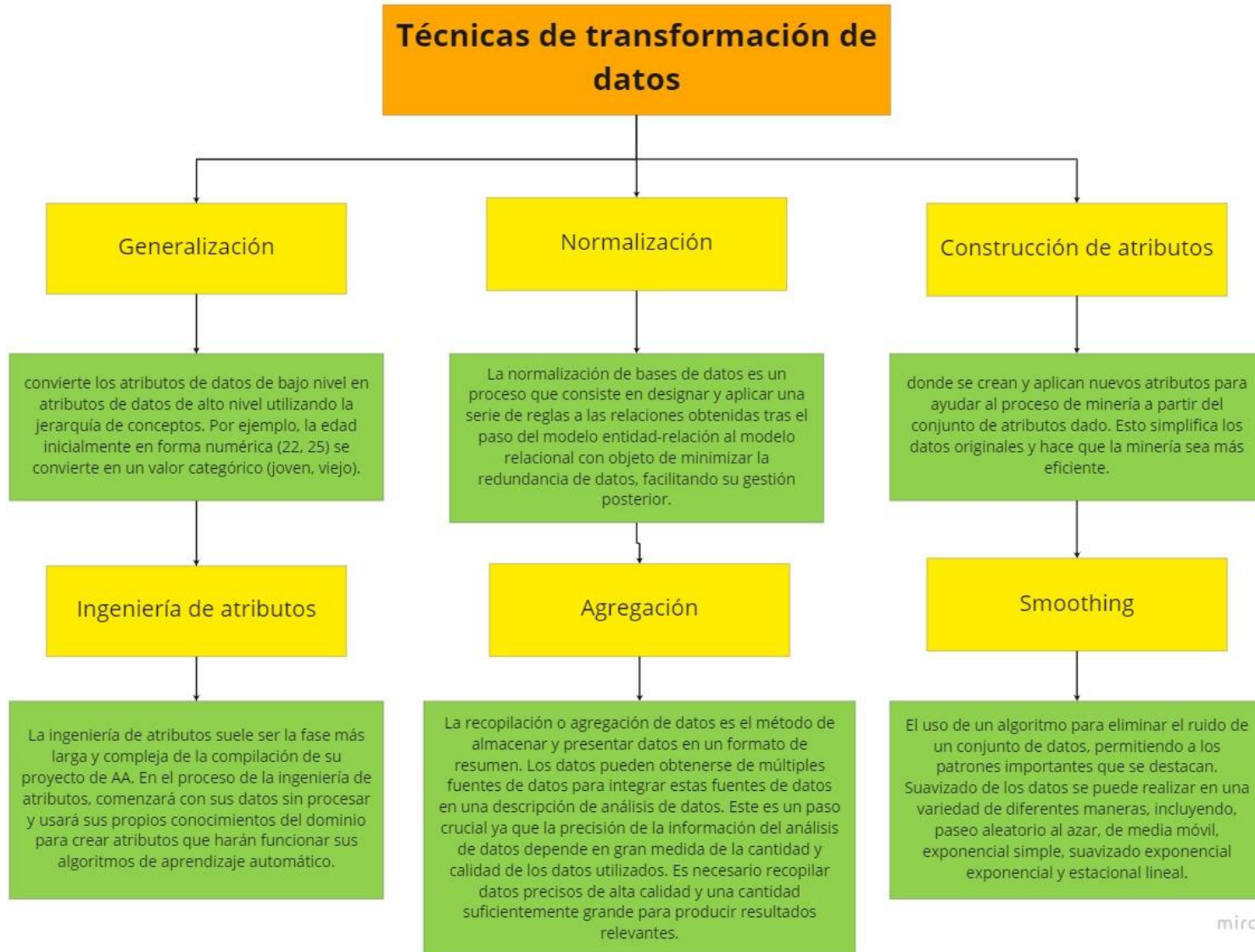
Identificar repositorios de datos y herramientas de generación de datos.



Identificar las técnicas de extracción de datos.

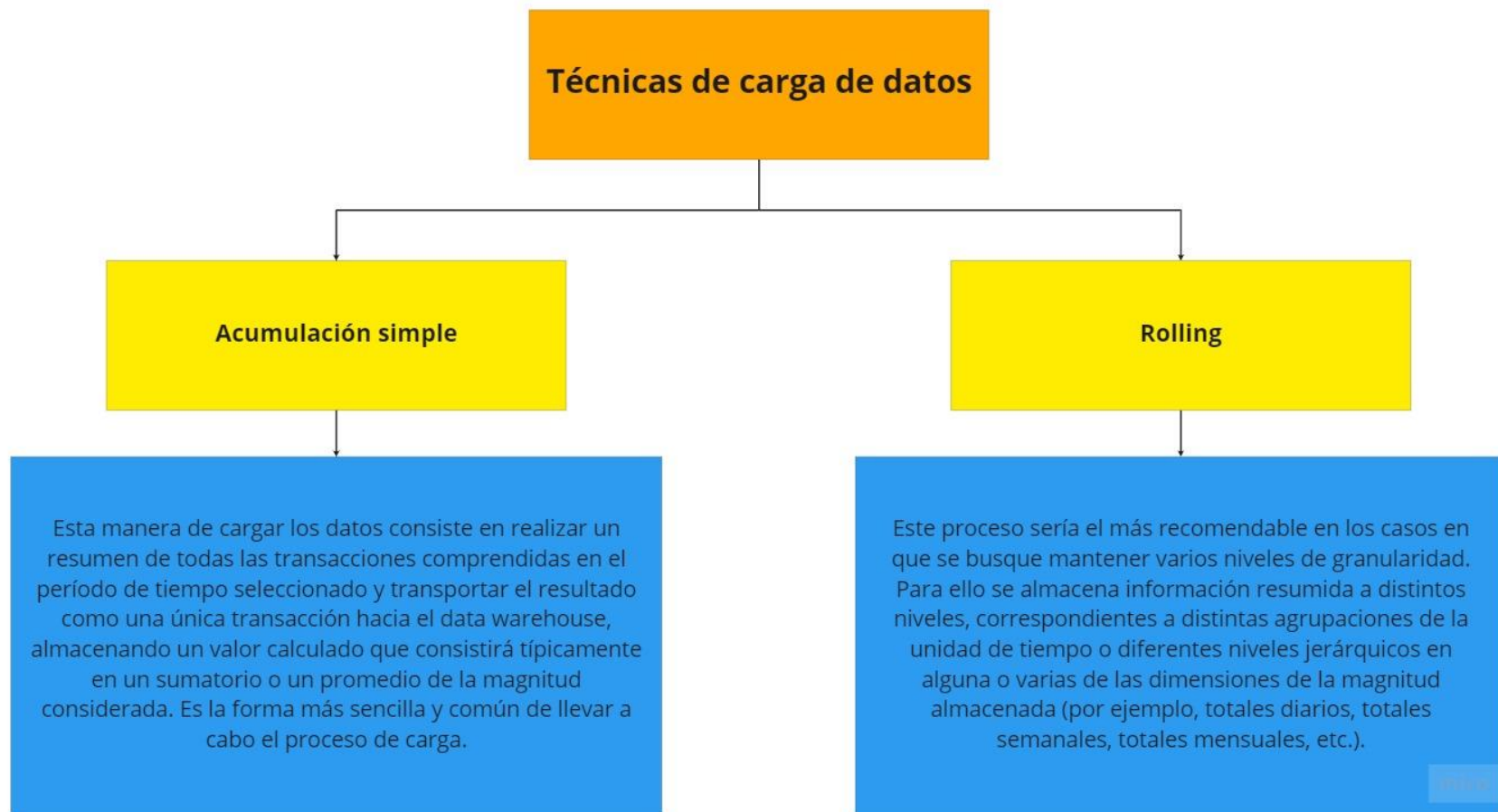


Identificar las técnicas de transformación de datos.



miro

Identificar las técnicas de carga de datos.



Conclusiones

Finalmente se concluyó que se analizaron algunos de los métodos recomendados para transformar datos. Esta es una tarea que puede llevar mucho tiempo, de hecho, se estima que podría consumir el 70% del tiempo total de un proyecto de Data Mining, pero las ventajas son enormes.

Una vez que los datos han sido preparados las relaciones que contienen quedan mejor expuestas, facilitando de esta manera el trabajo que tiene que hacer la herramienta que vaya a crear el modelo.

Bibliografía

Conecta Software. (2020, April 21). Tipos de fuentes de datos. Ficheros / Eventos.

Conecta Software. <https://conectasoftware.com/analytics/tipos-de-fuentes-de-ficheros/>

Datos de prueba. (2013, May 28). Globe Testing. <https://ahorasomos.izertis.com/globetesting/datos-de-prueba/>

Datos smoothing. (n.d.). Wikifx.Info. Retrieved October 2, 2022, from <http://www.wikifx.info/sp/datos-smoothing/>

Datos subjetivos y objetivos en enfermería: Diferencias clave. (2022, September 26). Historiadelaempresa.com. <https://historiadelaempresa.com/datos-subjetivos-vs-objetivos>

de los Santos, P. R. (2022, January 24). Datos de entrenamiento vs datos de test. Think Big. <https://empresas.blogthinkbig.com/datos-entrenamiento-vs-datos-de-test/>

Diferencia entre datos estructurados y no estructurados. (n.d.). Kyoceradocumentsolutions.Es. Retrieved October 2, 2022, from <https://www.kyoceradocumentsolutions.es/es/smarter-workspaces/insights-hub/articles/diferencia-entre-datos-estructurados-y-no-estructurados.html>

López, E. Á. S.-J. (2011, April). Tipos de datos. Unican.es. https://personales.unican.es/alvareze/OCW_CALC/3 CONTENIDOS/tema_3/t3_0_2.htm

Masters, T. y. (2021, March 30). Análisis de datos cuantitativos. Tesis y Masters Argentina. <https://tesisymasters.com.ar/analisis-de-datos-cuantitativos/>

Naeem, T. (2020a, June 23). ¿Qué es la extracción de bases de datos y por qué la necesita? Astera. <https://www.astera.com/es/type/blog/database-extraction/>

Naeem, T. (2020b, November 1). Comprender los datos estructurados, semiestructurados y no estructurados. Astera. <https://www.astera.com/es/type/blog/structured-semi-structured-and-unstructured-data/>

Naeem, T. (2020c, November 7). Repositorio de datos: importancia, desafíos y mejores prácticas. Astera. <https://www.astera.com/es/type/blog/data-repository/>

Preprocesar y normalizar datos, 4 pasos para limpiar y mejorar datos. (n.d.). Powerdata.es. Retrieved October 2, 2022, from <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/preprocesar-y-normalizar-datos-4-pasos-para-limpiar-y-mejorar-datos>

Procesos ETL: Extracción, Transformación, Carga. (n.d.). Powerdata.es. Retrieved October 2, 2022, from <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/288859/procesos-etl-extracci-n-transformaci-n-carga>

¿Qué es un almacén de datos? (n.d.). Oracle.com. Retrieved October 2, 2022, from <https://www.oracle.com/mx/database/what-is-a-data-warehouse/>

Sierra, Y. (2019, October 23). Análisis de datos estadísticos en Excel: Técnicas, Etapas e Importancia. #ADNCLOUD; Mediacloud. <https://blog.mdcloud.es/analisis-de-datos-estadisticos-que-es-y-como-aplicarlo-sin-esfuerzo/>

Transformación de datos en minería de datos – Acervo Lima. (n.d.). Acervolima.com. Retrieved October 2, 2022, from <https://es.acervolima.com/transformacion-de-datos-en-mineria-de-datos/>

Valbuena, D. F. (2020, April 3). Esquemas en Data Warehousing. Data Management. <https://datamanagement.es/2020/04/03/esquemas-data-warehousing/>

Velázquez, A. (2018a, September 17). ¿Qué son los datos ordinales? QuestionPro. <https://www.questionpro.com/blog/es/datos-ordinales/>

Velázquez, A. (2018b, September 24). ¿Qué son los datos nominales? QuestionPro. <https://www.questionpro.com/blog/es/datos-nominales/>

Wikipedia contributors. (n.d.). Datos binarios. Wikipedia, The Free Encyclopedia. https://es.wikipedia.org/w/index.php?title=Datos_binarios&oldid=133624562