**Analítica de Grandes Datos**

**Departamento de Ciencias de la Computación y la Decisión**

**Facultad de Minas**

**Universidad Nacional de Colombia**

Trabajo Nociones de Arquitectura de la Información

Versión: 2021-05-23

**responsables**

|  |
| --- |
| Nombre Completo – Documento de Identificación |
| 1. Daniel Múnera Arango 1.018.453.614 |
| 2. Cristian Camilo Londoño 1.128.270.374 |
| 3. Juan Mauricio Herrera 1.128.453.482 |
| 4. Juan David Ramirez Sanchez 1.128.430.332 |
|  |
| **REPO EN GITHUB: https://github.com/crlondono/AGD\_TrabajoModeloEntidadRelacion** |

1. **comprensión del negocio**
   1. **descripción del contexto del negocio.**

Debido al peligro inminente por un nuevo virus llamado SARS-COV2, en febrero 2020 se temía una posible futura propagación, el gobierno nacional de Colombia declaró estado de pandemia en el país, por lo cual, a partir de marzo de dicho año se impusieron una serie de medidas para contener la propagación del virus, anunciando la aplicación de un esquema de aislamiento preventivo obligatorio en todo el país, el cual incluyó entre otras medidas, confinamiento total obligatorio, cierre de establecimientos comerciales, limitación del transporte público de pasajeros por diversos medios, con la excepción de actividades esenciales como abastecimiento de productos de hogar y la continuidad de servicios médicos y hospitalarios. Teniendo en cuenta el gran cambio que tuvo que adoptar la sociedad en las diversas ciudades del país para asumir estas nuevas medidas impuestas por el gobierno, este trabajo se centra en el análisis del impacto de la pandemia en algunas de las dinámicas del componente de seguridad y convivencia ciudadana de Medellín. El análisis usa datos administrativos de denuncias policiales relacionados con temas de hurto a personas, violencia intrafamiliar, extorsión y hurto a residencias presentados en la ciudad.

* 1. **identificación del problema:**

Es necesario identificar los cambios que han tenido algunos ámbitos de la componente de seguridad y convivencia ciudadana en Medellín tras la adopción de las medidas de aislamiento, debido a que el comportamiento de esta puede haber cambiado y por ende se deban cambiar las medidas de seguridad y convivencia que se han implementado constantemente desde antes de la pandemia. Para lograr esto se requiere desarrollar un modelo de bases de datos que permita relacionar y hacer análisis sobre la afectación de la pandemia sobre estos componentes antes mencionados. Las restricciones que se pueden encontrar para realizar este análisis son la poca disponibilidad y actualidad de los datos, por lo cual se debe hacer una búsqueda detallada entre las diferentes fuentes para que la información obtenida sea verídica y confiable. Los beneficios obtenidos de este análisis, será el tener un panorama claro de los indicadores de seguridad y convivencia en la ciudad de Medellín, y así proveer evidencia para ayudar en la toma de decisiones de política pública y el diseño y/o ajuste de estrategias policiales en un escenario de pandemia.

* 1. **determinación de objetivos:**

Mediante los resultados de este análisis se podrá encontrar nuevas tendencias y comportamientos de la componente de seguridad y convivencia ciudadana en Medellín, por lo cual se tendrán ventajas en términos de costos y tiempos referente a la toma de decisiones para la implementación de nuevas estrategias para la mejora de esta problemática. Por lo cual, para lograr esto, se plantea desarrollar los siguientes objetivos:

1. Desarrollar un modelo entidad - relación para relacionar los datos de propagación del Covid-19 con datos de hurto a personas, violencia intrafamiliar, extorsión y hurto a residencia en Medellín.
2. Realizar un análisis de los cambios en la información comparando el comportamiento de la seguridad y convivencia ciudadana antes y después de la pandemia.
3. Mejorar la comprensión sobre el impacto de las emergencias sanitarias en las dinámicas de la seguridad y convivencia ciudadana de Medellín.
   1. **evaluación de la situación actual:**

La pandemia tuvo un impacto en las estrategias policiales, ya que estos tuvieron que ajustar diversos aspectos de su operatividad para incorporar responsabilidades adicionales a su labor habitual de mantener el orden público. Los entes policiales respondieron a la emergencia sanitaria ejerciendo, por ejemplo, el control del cumplimiento a las restricciones de movilidad o a la clausura de sectores económicos. Por esto, la policía también necesitó ajustarse para frenar los contagios en sus filas, en ocasiones reduciendo el personal disponible por determinados períodos de tiempo, lo cual ocasiona que haya menor cuerpo policial disponible para atender los eventos de criminalidad presentados.

Según [1], “hubo un aumento del 14 por ciento en los casos de violencia intrafamiliar y del 18 por ciento en robos de motos en las últimas dos semanas. [Sin embargo, hubo disminución en los homicidios y hurtos a personas](https://www.elcolombiano.com/antioquia/seguridad/el-crimen-en-tiempos-de-covid-19-HM12692656)”.

Debido a esto es necesario identificar el comportamiento actual de los sectores de Medellín para plantear nuevas estrategias de seguridad y convivencia ciudadana.

1. **comprensión de los datos**
   1. **recolección de datos**

Para la recolección de datos, se emplearon dos fuentes, MEData que es el portal donde se encuentran los datos abiertos generados y publicados por las diferentes dependencias de la Alcaldía de Medellín. Y GeoMedellin, que es el portal donde se encuentra el resultado de los procesos de gestión y uso de tecnologías de información, para que datos e información estratégica se dispongan de forma fácil, asequible y gratuita.

**Bases de datos obtenidas**

|  |  |  |
| --- | --- | --- |
| Nombre | Descripción | Fuente |
| COVID19 Casos Confirmados Medellín | Información asociada al centroide del municipio, a partir del consolidado de casos por comunas y corregimientos de la ciudad de Medellín. | <https://geomedellin-m-medellin.opendata.arcgis.com/datasets/covid19-casos-confirmados-opendata/data?geometry=-75.824%2C6.192%2C-75.361%2C6.311> |
| Comunas y Barrios Medellín | Representa el Limite de comuna o corregimiento de acuerdo a la conformación de predios de la ciudad.  Esta base de datos fue construida a través de la investigación en las siguientes fuentes presentadas | https://www.medellin.gov.co/normograma/docs/astrea/docs/R\_CATASTRO\_0001\_2005.htm  https://www.medellin.gov.co/normograma/docs/astrea/docs/R\_SHMED\_0001\_2010.htm |
| Extorsión | Hechos relacionados con la extorsión en la ciudad de Medellín | <http://medata.gov.co/dataset/extorsi%C3%B3n> |
| Hurto a residencia | Hechos relacionados con el hurto a residencia en la ciudad de Medellín | <http://medata.gov.co/dataset/hurto-residencia> |
| Reincidencia violencia intrafamiliar | Hechos relacionados con violencia intrafamiliar en la ciudad de Medellín | <http://medata.gov.co/dataset/reincidencia-violencia-intrafamiliar> |
| Hurto a persona | Hechos relacionados con Hurto a persona en la ciudad de Medellín | <http://medata.gov.co/dataset/hurto-persona> |

Las dificultades y soluciones presentadas en la obtención de los datos fueron las siguientes:

**Bases de datos desactualizadas:** se descargó diversas bases de datos, y se hizo una validación de los datos referente a la fecha de los registros para determinar su actualidad, debido a que las descripciones previas de los portales no eran 100% verdaderas respecto al tema.

**Links de accesos desactualizados:** en muchas ocasiones los links para descargar las bases de datos se encontraban caídas o desactualizadas. Para las fuentes de datos que finalmente se usaron en este trabajo se validó la viabilidad de acceso a los datos de manera recurrente.

**Más de una fuente del mismo tema**: se descargaron las bases de datos encontradas de un mismo tema y se compararon sus respectivos campos y calidad para determinar cuál es la más viable para el análisis.

**Baja calidad en algunos registros y campos:** se debió realizar una depuración de los datos debido a que había registros con campos incompletos y las bases de datos tenían gran cantidad de campos que no eran necesarios para este estudio, por lo cual se depuró mediante la implementación y uso de los siguientes scripts python:

* Organizar\_Archivo\_Covid.py
* Organizar\_Archivo\_Extorsion.py
* Organizar\_Archivo\_Hurto\_Persona.py
* Organizar\_Archivo\_Hurto\_Residencia.py
* Organizar\_Archivo\_Violencia\_Intrafamiliar.py

**Datos sin normalizar y consolidación:** se debió normalizar algunos campos de los datos, y realizar una consolidación de las bases de datos de extorsión, hurto\_a\_persona, hurto\_a\_residencia y reincidencia\_violencia en la tabla hechos\_delictivos para un mejor manejo de esta, esto gracias a que compartían gran variedad de campos de interés, por lo cual se realizó mediante la implementación y uso de los siguientes scripts python:

* Normalizar\_Comunas\_Barrios\_Hechos\_delictivos.py
* Normalizar\_Comunas\_Casos\_Covid.py
* Normalizar\_Hechos\_Delictivos.py
  1. **descripción de datos (diccionario):**

En base al procesamiento previo realizado a las bases de datos obtenidas, se construyeron las siguientes tablas para la elaboración del modelo de entidad relación.

|  |  |  |
| --- | --- | --- |
| Casos\_Covid\_Medellin | | |
| Nombre del atributo / variable | Formato o Tipo de Dato | Descripción |
| ID | numerico | Código único de registro |
| Codigo\_Sexo | NVarchar(1) | Sexo del paciente |
| Codigo\_Comuna | numerico | Código identificación comuna |
| Edad | numerico | Edad del paciente |
| Fecha\_Inicio\_Caso | Date | Fecha en que inició a presentar los síntomas |
| Fecha\_Fin\_Caso | Date | Fecha finalización de síntomas, construido a partir de la referencia [2] donde concluyen que el tiempo promedio de la enfermedad es de 12 días |

|  |  |  |
| --- | --- | --- |
| Hechos\_Delictivos | | |
| Nombre del atributo / variable | Formato o Tipo de Dato | Descripción |
| ID | Int | Código único de registro |
| Fecha\_Hecho | Date | Fecha en que ocurrió el hecho |
| Tipo\_Hecho | numerico | Tipo acto delictivo que sufrió la víctima |
| Codigo\_Sexo | NVarchar(1) | Sexo de la victima |
| Edad | Int | Edad de la víctima |
| Modalidad | NVarchar(30) | Forma como se materializa el hecho |
| Ocupacion | NVarchar(30) | Ocupación de la víctima o del presunto indiciado, Mirar en contexto con conducta |
| Codigo\_Comuna | numerico | Código de la comuna donde ocurrieron los hechos |
| Codigo\_Barrio | numerico | Código del barrio donde ocurrieron los hechos |
| Bien | NVarchar(3) | Nombre del bien |

|  |  |  |
| --- | --- | --- |
| Comunas\_Medellin | | |
| Nombre del atributo / variable | Formato o Tipo de Dato | Descripción |
| Codigo\_Comuna | numerico | Código de la comuna de medellín |
| Nombre\_Comuna | Nvarchar(30) | Nombre de la comuna de medellín |

|  |  |  |
| --- | --- | --- |
| Barrios\_Medellin | | |
| Nombre del atributo / variable | Formato o Tipo de Dato | Descripción |
| Código\_Comuna | numerico | Código de la comuna de medellín |
| Codigo\_Barrio | numerico | Código del barrio de medellín |
| Nombre\_Barrio | NVarchar(30) | Nombre del barrio de medellín |

|  |  |  |
| --- | --- | --- |
| Sexos | | |
| Nombre del atributo / variable | Formato o Tipo de Dato | Descripción |
| Codigo\_Sexo | NVarchar(1) | Codigo de sexo |
| Descripcion | NVarcahr(20) | Descripcion identificacion de sexo |

|  |  |  |
| --- | --- | --- |
| Dimension\_Tiempo | | |
| Nombre del atributo / variable | Formato o Tipo de Dato | Descripción |
| Fecha | Date | Fecha en formato de AAAA-MM-DD |
| Dia | numerico | Dia en formato de día semana de 1 a 7 |
| Mes | numerico | Mes en formato numérico de 1 a 12 |
| Anio | numerico | Año en formato de 4 dígitos |

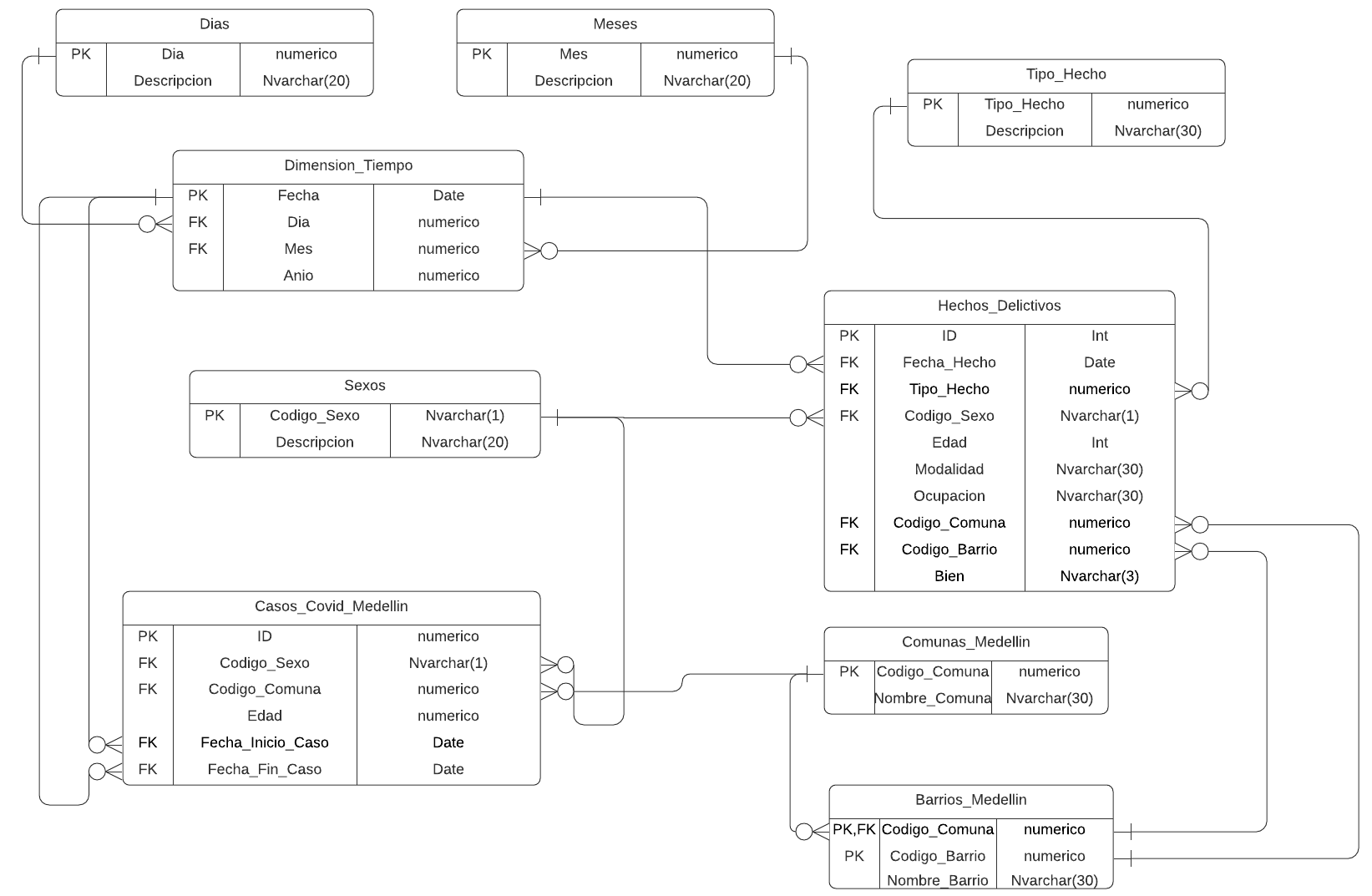
|  |  |  |
| --- | --- | --- |
| Días | | |
| Nombre del atributo / variable | Formato o Tipo de Dato | Descripción |
| Dia | numerico | Dia en formato de día semana de 1 a 7 |
| Descripcion | numerico | Nombre del dia de la semana |

|  |  |  |
| --- | --- | --- |
| Meses | | |
| Nombre del atributo / variable | Formato o Tipo de Dato | Descripción |
| Mes | numerico | Mes en formato numérico de 1 a 12 |
| Descripcion | numerico | Nombre del mes |

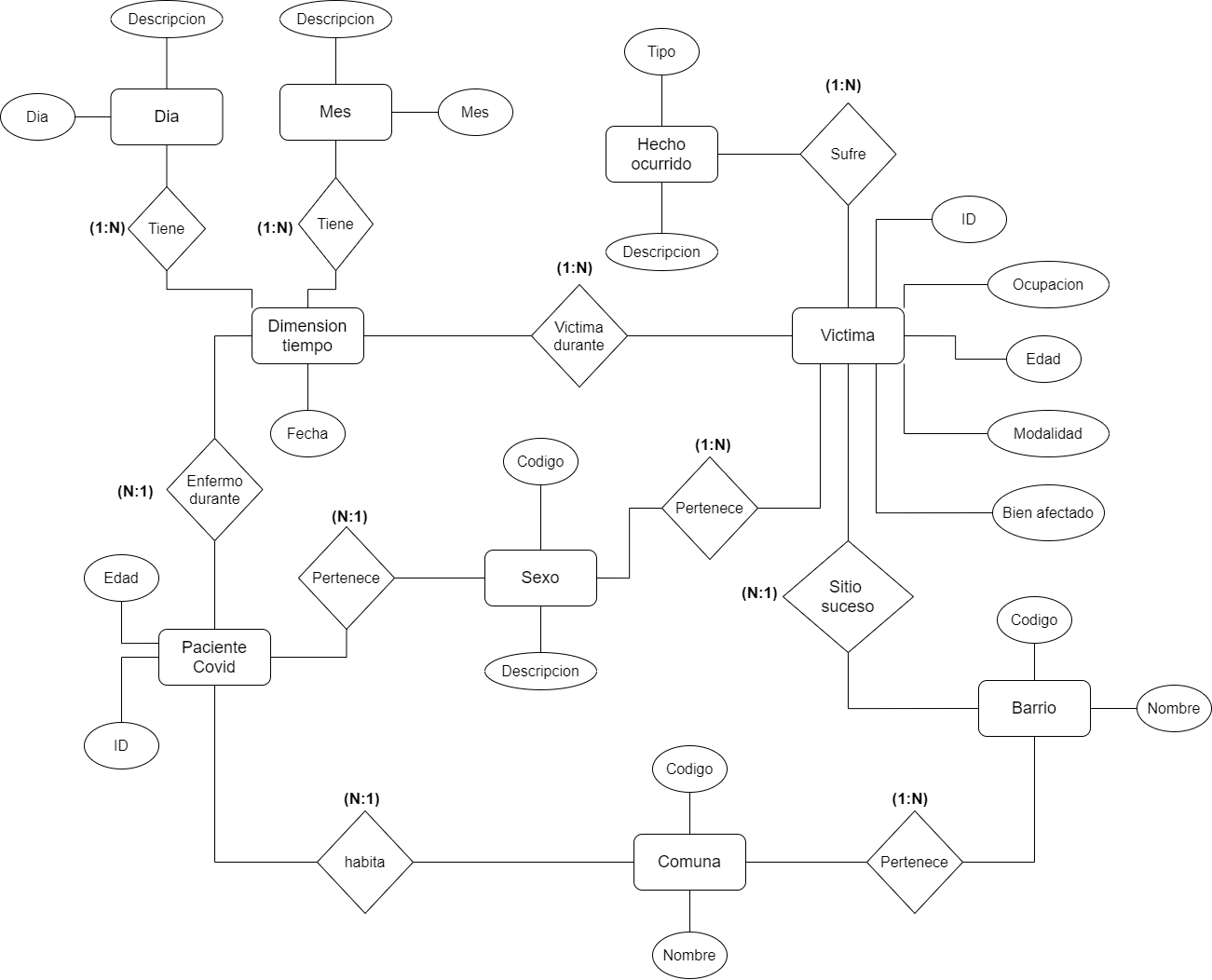
|  |  |  |
| --- | --- | --- |
| Tipo\_Hecho | | |
| Nombre del atributo / variable | Formato o Tipo de Dato | Descripción |
| Tipo\_Hecho | Date | Código numérico de 1 a 4 para clasificación de hecho |
| Descripcion | NVarchar(30) | Nombre del hecho |

* 1. **modelo del domino**

***Observación****: Incluye el gráfico del modelo del dominio que representa la estructura de datos.*

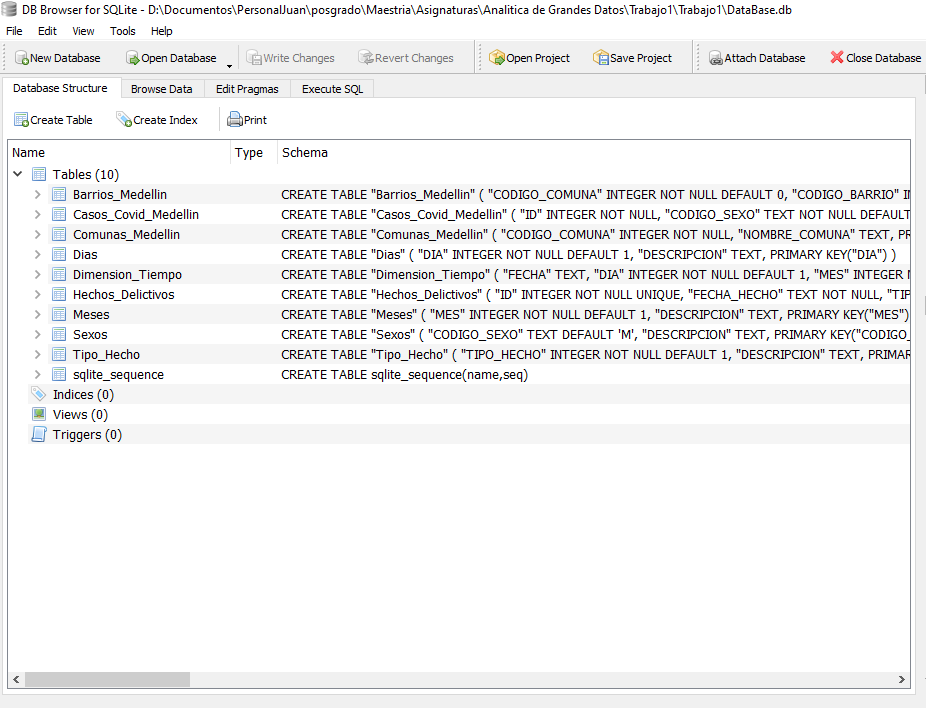
**

1. **modelo entidad-relación**
   1. **toma de pantalla del modelo E-R**

**

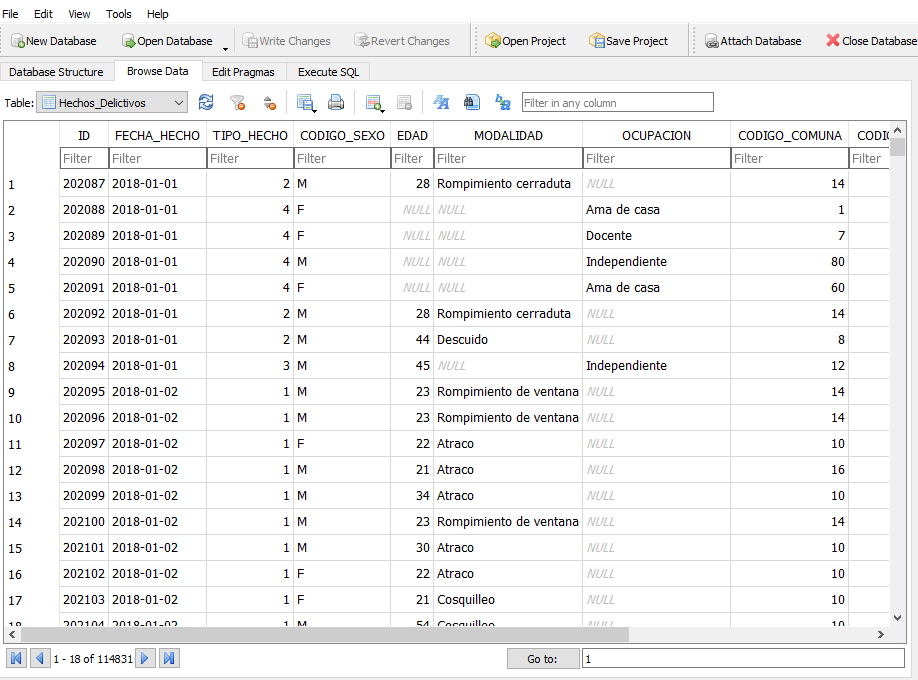
* 1. **sentencia o consulta de creación del tabla(s)**

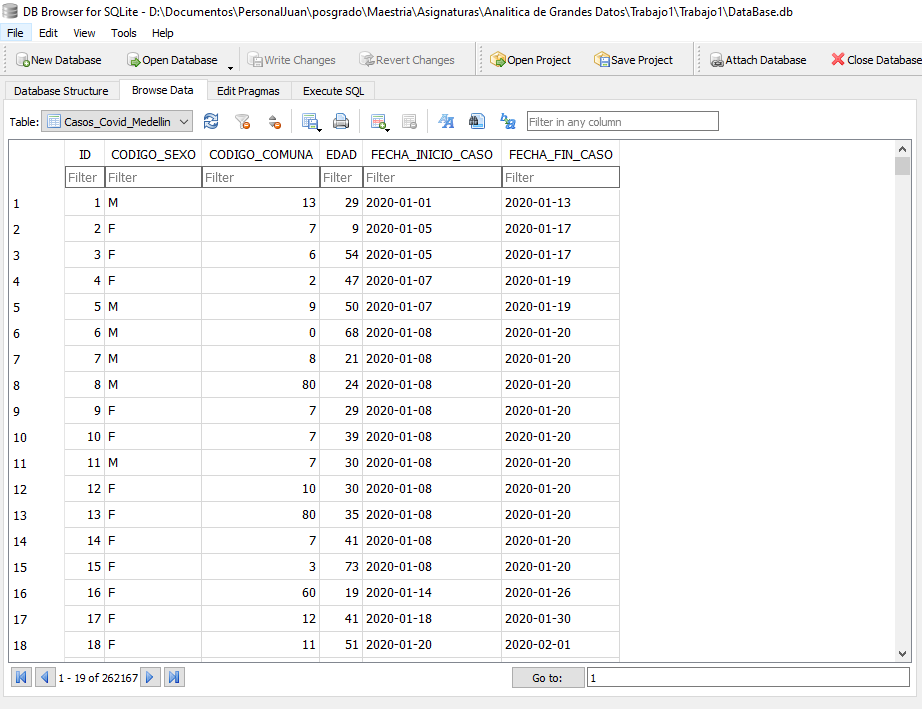
Las sentencias ejecutadas se encuentran en el repositorio GIT adjunto. A continuation una imagen de las tablas construidas.

******

* 1. **sentencias para insertar datos**

***Observación****:* Las sentencias ejecutadas se encuentran en el repositorio GIT adjunto. A continuación una imagen de la data cargada.

**

**

* 1. **sentencias de consulta**

***Observación****:*

Las consultas realizadas se enuncian a continuación, y sus respectivos screenshot y querys se encuentran en los archivos “screenshot\_QuerySQL\_T1.3.4.Consultar\_Datos” y “T1.3.4.Consultar\_Datos.sql”

1. **Hechos delictivos por comuna.**

1.1 número de tipo de hecho por comuna 2018.

1.2 número de tipo de hecho por comuna 2019.

1.3 número de tipo de hecho por comuna 2020.

1.4 Media de casos por comuna 2020

1.5 promedio hechos por comuna 2019

1.6 promedio hechos por comuna 2020

1.7 mínimo de hecho por comuna 2018.

1.8 mínimo de hecho por comuna 2019.

1.9 mínimo de hecho por comuna 2020.

1.10 top 5 de hecho por comuna 2018.

1.11 top 5 de hecho por comuna 2019.

1.10 top 5 de hecho por comuna 2020.

1. **Hechos delictivos por barrio.**

2.1 Número de tipo de hecho por barrio 2018.

2.2 número de tipo de hecho por barrio 2019.

2.3 Número de tipo de hecho por barrio 2020.

2.4 mínimo de hecho por barrio 2018.

2.5 mínimo de hecho por barrio 2019.

2.6 mínimo de hecho por barrio 2020.

2.7 top 5 de hecho por barrio 2018.

2.8 top 5 de hecho por barrio 2019.

2.9 top 5 de hecho por barrio 2020.

1. **Hechos delictivos por sexo.**

3.1 número de tipo de hecho por sexo 2018.

3.2 número de tipo de hecho por sexo 2019.

3.3 número de tipo de hecho por sexo 2020.

3.4 minimo de tipo de hecho por sexo 2018.

3.5 minimo de tipo de hecho por sexo 2019.

3.6 minimo de tipo de hecho por sexo 2020.

3.7 top 5 de tipo de hecho por sexo 2018.

3.8 top 5 de tipo de hecho por sexo 2019.

3.9 top 5 de tipo de hecho por sexo 2020.

1. **Hechos delictivos por rango de edad.**

4.1 número de tipo de hechos por rangos de edad (0 - 20, 21 - 40, 41 - 60, 60 - 120) 2018.

4.2 número de tipo de hechos por rangos de edad (0 - 20, 21 - 40, 41 - 60, 60 - 120) 2019.

4.3 número de tipo de hechos por rangos de edad (0 - 20, 21 - 40, 41 - 60, 60 - 120) 2020.

4.4 minimo de de tipo de hechos por rangos de edad 2018.

4.5 minimo de de tipo de hechos por rangos de edad 2019.

4.6 minimo de tipo de hechos por rangos de edad 2020.

4.7 top 5 de tipo de hechos por rangos de edad 2018.

4.8 top 5 de tipo de hechos por rangos de edad 2019.

4.8 top 5 de tipo de hechos por rangos de edad 2020.

1. **Hechos delictivos por ocupación.**

5.1 número de tipo de hecho por ocupación 2018.

5.2 número de tipo de hecho por ocupación 2019.

5.3 número de tipo de hecho por ocupación 2020.

5.4 mínimo de tipo de hecho por ocupación 2018.

5.5 mínimo de tipo de hecho por ocupación 2019.

5.6 mínimo de tipo de hecho por ocupación 2020.

5.7 top 5 de tipo de hecho por ocupación 2018.

5.8 top 5 de tipo de hecho por ocupación 2019.

5.9 top 5 de tipo de hecho por ocupación 2020.

1. **Hechos delictivos por tipo de bien.**

6.1 número de tipo de bien por tipo de hecho por comuna 2018.

6.2 número de tipo de bien por tipo de hecho por comuna 2019.

6.3 número de tipo de bien por tipo de hecho por comuna 2020.

6.4 mínimo de tipo de bien por tipo de hecho por comuna 2018.

6.5 mínimo de tipo de bien por tipo de hecho por comuna 2019.

6.6 mínimo de tipo de bien por tipo de hecho por comuna 2020.

6.7 top 5 de tipo de bien por tipo de hecho por comuna 2018.

6.8 top 5 de tipo de bien por tipo de hecho por comuna 2019.

6.9 top 5 de tipo de bien por tipo de hecho por comuna 2020.

1. **Casos covid por comuna.**

7.1 número de casos covid por comuna 2020.

7.2 mínimo de casos covid por comuna 2020.

7.3 top 5 de casos covid por comuna 2020.

1. **Casos covid por sexo.**

8.1 número de casos covid por sexo 2020.

8.2 minimo de casos covid por sexo 2020.

8.3 top 5 de casos covid por sexo 2020.

1. **Casos covid por rangos de edad.**

9.1 número de casos covid por rangos de edad (0 - 20, 21 - 40, 41 - 60, 60 - 120)

9.2 mínimo de casos covid por rangos de edad 2020.

9.3 top 5 de casos covid por rangos de edad 2020.

1. **Hechos delictivos por comuna por mes.**

10.1 número de hechos delictivos por comuna por mes 2018

10.2 número de hechos delictivos por comuna por mes 2019

10.3 número de hechos delictivos por comuna por mes 2020

10.4 mínimo de hechos delictivos por comuna por mes 2018.

10.5 mínimo de hechos delictivos por comuna por mes 2019.

10.6 mínimo de hechos delictivos por comuna por mes 2020.

10.7 top 5 hechos delictivos por comuna por mes 2018.

10.8 top 5 hechos delictivos por comuna por mes 2019.

10.9 top 5 hechos delictivos por comuna por mes 2020.

1. **Casos covid por comuna por mes.**

11.1 número de casos covid por comuna por mes 2020

11.2 mínimo de casos covid por comuna por mes 2020.

11.3 top 5 de casos covid por comuna por mes 2020.

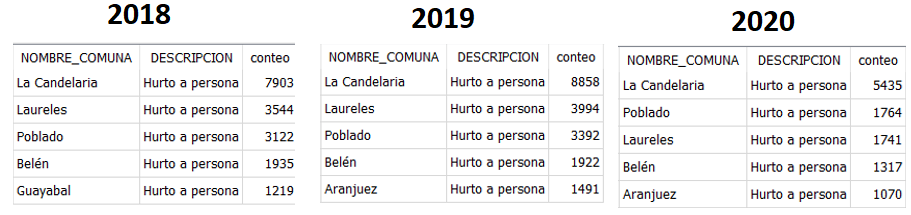
**Análisis de consultas:**

En base a las consultas realizadas se tienen el siguiente análisis:

1. **Hechos delictivos por comuna.**

La menor actividad delictiva en el 2018 fue de Extorsión en la comuna de Santa Elena, teniendo en cuenta la actividad de pandemia en 2020 la actividad menos presentada en este año fue de hurto a residencia en la comuna de San Sebastián de Palmitas, esto debido a que en pandemia muchas residencias de descanso permanecieron solas por largos periodos.

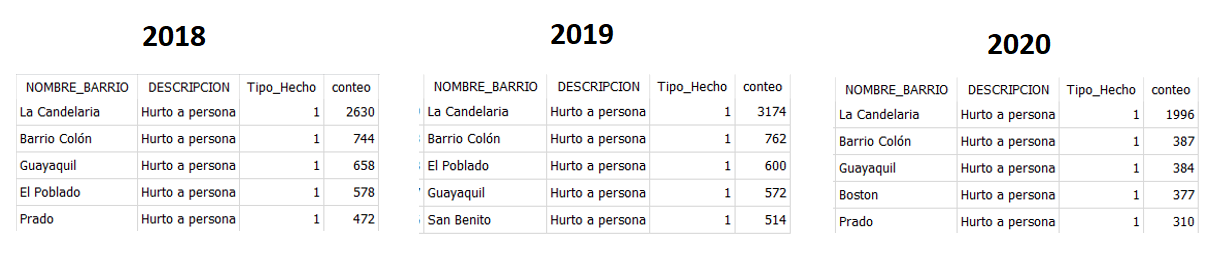
En los últimos 3 años la comuna La Candelaria presenta la mayor cantidad de actividad delictiva, pero analizando el número de casos en época de pandemia estos se redujeron de forma considerable, lo cual hace razonable pensar que fue debido a la cuarentena obligatoria teniendo en cuenta que el mayor hecho presentado fue hurto a persona.



1. **Hechos delictivos por barrio.**

La menor actividad delictiva presentada en el 2018 a 2020 no varió respecto al barrio, ya que se presentó en Santo Domingo con solo una variación en la actividad delictiva referente a violencia intrafamiliar y extorsión siendo esta última en época de pandemia.

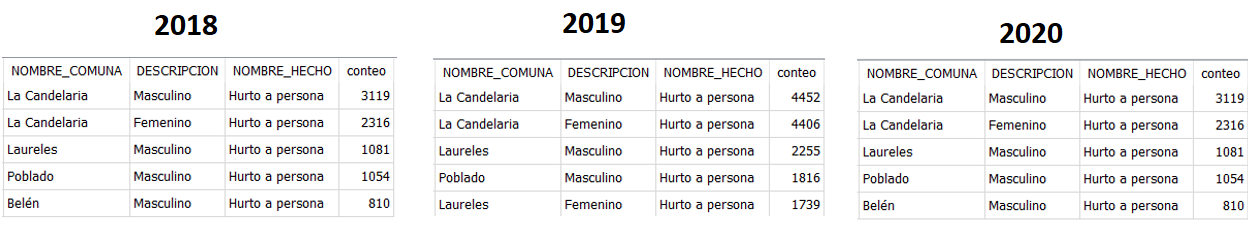
Respecto a los barrios, el barrio la candelaria, Colón y Guayaquil siguen liderando con la mayor cantidad de actividad delictiva, el estado de pandemia solo genera notoria diferencia en las actividades delictivas del barrio El Poblado, ya que en época de pandemia esto no se presentó con tanta frecuencia como en años anteriores.



1. **Hechos delictivos por sexo.**

La menor actividad delictiva fue en San Sebastián de Palmitas a una persona de sexo femenino bajo la modalidad de hurto a persona.

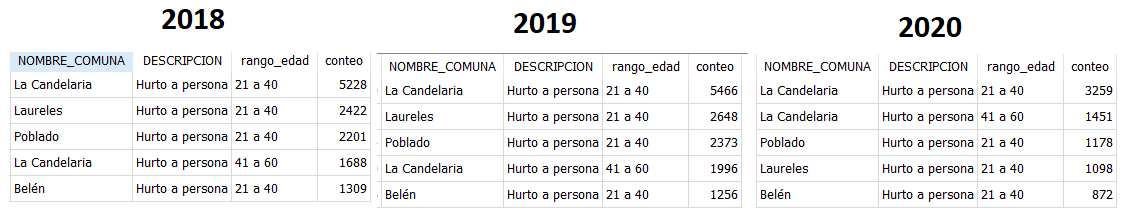
Se evidencia que a través de los años el sexo masculino presenta mayor afectación por actividades delictivas que el sexo femenino.



1. **Hechos delictivos por rango de edad.**

Se evidencio una constante de un mínimo de hechos delictivos por rango de edad de 0 a 20 años bajo la modalidad de Extorsión en barrios como Castilla, Aranjuez y San Javier.

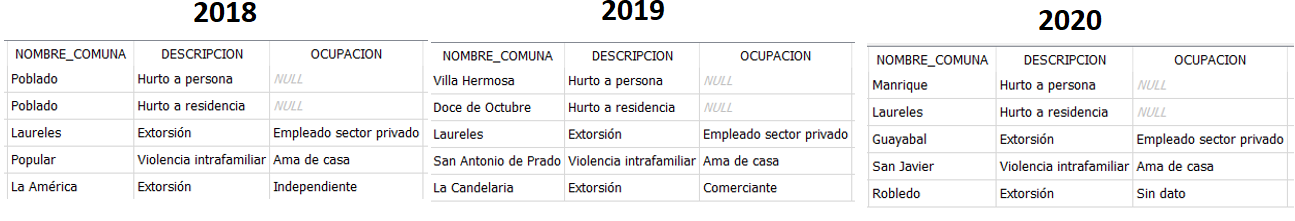
También se mantuvo como mayor afectado a hurto a las personas con edad de entre 0 a 20 años



1. **Hechos delictivos por ocupación.**

La menor actividad delictiva en época de pandemia se presentó en la comuna de Robledo y fue a una persona que ejerce la ocupación de administrador siendo afectado bajo la modalidad de extorsión.

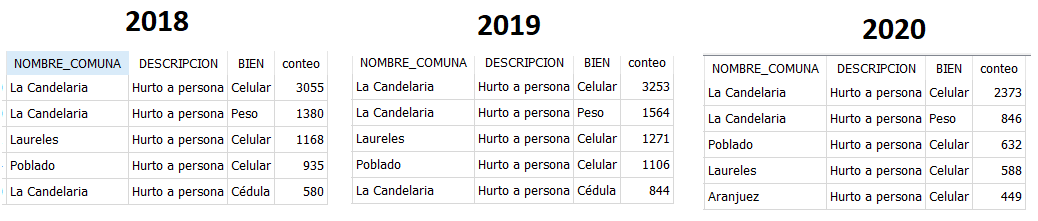
Se evidenció que se mantuvo una constante en los años de afectaciones a personas con ocupación de empleados del sector privado y amas de casa, teniendo un aumento este último en la modalidad de violencia intrafamiliar, lo cual es coherente referente a la actividad de pandemia donde los ciudadanos se encontraban el mayor tiempo en sus domicilios.



1. **Hechos delictivos por tipo de bien.**

El mínimo bien robado fueron objetos de bajo costo como Alambre o licores, lo cual no fueron objetos que se presentaron en época de pandemia, y es coherente referente a que las discotecas se encontraban limitadas.

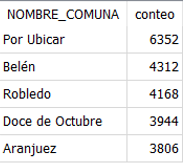
Se pudo visualizar que el bien más robado son el celular y el peso o dinero, aunque se ve una disminución importante en el sector de El Poblado referente a este bien, esto debido a las acciones de pandemia donde había menos circulación de los sectores urbanos.



1. **Casos covid por comuna.**

La menor cantidad de contagios presentada fue en San Sebastián de Palmitas, misma comuna donde se presentaron menor cantidad de hechos delictivos.

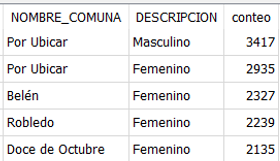
La mayor cantidad de casos presentados fue en barrios de mayor concurrencia rural lo cual coincide con los barrios donde la actividad delictiva referente a hurto persona se presenta con cantidades significativas.



1. **Casos covid por sexo.**

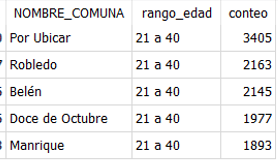
El mínimo de contagiados coincide con San Sebastián de Palmitas y fue referente a una persona de sexo masculino.

La mayor cantidad de personas contagiadas es hombre, lo cual explica porque los hechos delictivos bajaron significativamente en época de pandemia debido a que la actividad delictiva es ejecutada con más frecuencia por personas con sexo masculino.



1. **Casos covid por rangos de edad.**

La menor cantidad de personas contagiadas fue de 0 a 20 años en San Sebastián de Palmitas. La mayor cantidad de contagiados en las comunas son personas de edad entre 21 a 40 años, que sería el rango de edad que comprende la mayor parte de la personas que laboran y se desplazan por las zonas y es acorde a la mayor cantidad de personas que sufren hechos delictivos como lo son hurto a persona.



1. **Hechos delictivos por comuna por mes.**

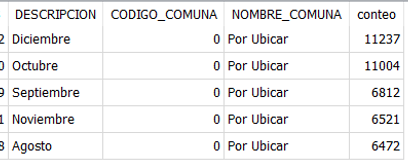
La menor cantidad de actividades delictivas se presentaron en el mes de enero en los barrios Popular y San Sebastián de Palmitas.

La mayor cantidad de Hechos delictivos se presentaron en el 2020 se presentaron en enero, a partir de febrero referente a los otros años, tuvo una disminución significativa, es posible que esto sea debido a que a pesar de que aún no se implementan acciones de pandemia, los ciudadanos ya estaban iniciando a adoptar medidas de prevención.



1. **Casos covid por comuna por mes.**

Es evidente que la mayor cantidad de casos presentó los últimos meses del año 2020 ya que el contagio se propaga generando infectados de forma exponencial



**Recomendación en base al análisis y conclusiones:**

En base a las consultas realizadas y analizadas previamente, se evidencia de forma general, que las actividades de hechos delictivos tuvieron una disminución significativa en las diversas comunas de Medellín, esto debido a que la mayor modalidad presentada en acto delictivo es referente a hurto persona, según la información consultada. Este cambio en cifras puede ser debido a que las actividades de pandemia redujeron la circulación de personas por los sectores más peligrosos de Medellín que se según las evidencias son la Candelaria, el Poblado y Laureles.

También se evidenció, según la información, que las actividades delictivas para 2020 a pesar de que los barrios más peligrosos disminuyeron su cifras delictivas, continúan encabezando el listado de sectores más peligrosos, pero a pesar de esto, es importante notar que en época de pandemia se presentó un aumento en el comportamiento de hechos delictivos en el sector de Aranjuez, lo cual nos puede dar a pensar que en este sector la ciudadanía no cumplió de igual medida las normas de cuarentena, generando una circulación significativa de personas en las calles, facilitando de esta forma el desarrollo de esto hechos delictivos.

1. **mongodb**
   1. **sentencias de creación del documento(s)**

***Observación****: Los códigos para la creación de los documentos se encuentran en el archivo T1.4.1.Creacion\_Documentos.zip en el repositorio del trabajo.*

* 1. **sentencias de consulta**

***Observación****: Las consultas referentes a la exploración básica de los datos en mongo, se encuentran en el archivos T1.4.1.Mongo\_db\_querys .zip y T1.4.2.Consultar\_Datos\_Mongo.docx en el repositorio del trabajo y son análogas a las realizadas en SQL.*

1. **análisis de lectura**

***Observación****: Considerando el artículo: “*The Definitive Guide to Graph Databases for the RDBMS Developer*” de Neo4J. Compartido en las carpeta de lecturas recomendadas. Analice y responda cada pregunta en máximo 150 palabras:*

1. **¿Cuáles son las limitaciones, que se pueden inferir de la lectura, para migrar los conjuntos de datos relacionales a NoSQL?**

Dentro de las desventajas que encontramos de una base de datos NoSQL, que puede llegar a ser las limitaciones por las cuales no se llevarían las bases de datos relacionales a NoSQL, encontramos:

* La mayoría de las bases NoSQL no contemplan la atomicidad de las instrucciones y la integridad de los datos.
* Los grandes problemas de compatibilidad entre instrucciones SQL. Las nuevas bases NoSQL utilizan sus propias características en el lenguaje de consulta y esto hace que no sean completamente compatibles con SQL.
* Una falta de estandarización. Existen diversidad de bases de datos NoSQL y no se encuentra aún un estándar como sí lo tiene las relacionales.

1. **¿Cuáles limitaciones adicionales que se deben considerar, a parte de las mencionadas en el artículo?**

Otras limitaciones adicionales que se deben tomar en cuenta son:

1. Los desarrolladores deben implementar su propio código, lo cual agrega mucha más complejidad al sistema.
2. Puede no haber una estandarización en el lenguaje para los diferentes ambientes de trabajo en las empresas, no se tiene un estándar definido entre los diferentes motores que ofrecen el servicio.

3. Las herramientas que se ofrecen para la administración de estas herramientas, suelen tener acceso por consola, al no tener una interfaz gráfica se requiere un conocimiento amplio de las instrucciones a utilizar para su mantenimiento.

1. **¿Cuáles son las razones (criterios) que se deben considerar para migrar un conjunto de datos relacionados a NoSQL?**

Las ventajas que encontramos en una base de datos Nosql son:

1. Son bases escalables y descentralizadas, lo cual permite soportar estructuras distribuidas.
2. Más flexibles y abiertas, permitiendo que se adapten a la necesidad de proyectos de manera más rápida que los modelos entidad-relación.
3. **Escalabilidad Horizontal:** Capaces (debido a la estructura distribuida) de crecer en número de máquinas en lugar de tener que estar en grandes volúmenes de almacenamiento.
4. Permiten ser ejecutadas en máquinas de pocos recursos.
5. Mucho más rápidas en consultas de grandes volúmenes de datos que las tradicionales entidad-relación.

Según esto, la forma para definir cuándo usar un conjunto de datos Nosql es según el problema que tengamos dentro de nuestros modelos de datos y nuestro contexto de negocio. Es recommendable usar Nosql cuando se dan los siguientes casos:

1. Análisis de grandes cantidades de datos
2. Captura y proceso de eventos en poca cantidad de tiempo.
3. Motores de inteligencia complejos, que necesitan ser procesados en la menor cantidad de tiempo posible. (tiempo real)

En general las bases de datos NoSQL, permiten la ejecución de procesos mucho más complejos con mayor volumen de datos.

1. **referencias**

[1] Diario AS Colombia, “Cuarentena en Medellín: ¿cómo ha afectado el índice de crimen?”, 28/04/2020. [ONLINE] https://colombia.as.com/colombia/2020/04/28/actualidad/1588095124\_721172.html

[2] BBC News Mundo, “Coronavirus: ¿cuándo una persona enferma de covid-19 deja de ser contagiosa (tenga o no síntomas)?”,2020, [ONLINE] <https://www.bbc.com/mundo/noticias-54828774>