UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

INGENIERA EN CIENCIAS DE LA COMPUTACIÓN

BASE DE DATOS AVANZADA

Tema: Catastro Turístico de Establecimientos a Nivel Nacional Informe Final - Proyecto Integrador

DOCENTE:

Ing. Eduardo Encalada

INTEGRANTES:

- Diana Bermeo (Paralelo B)
 - Carlos Castillo
 - Luis Granda
 - Rosa Palacios

PERIODO ACADÉMICO:

abril 2022 - agosto 2022

Tabla de contenido

1	Introducción	3
2	Repositorio GitHub	3
3	Datos base	4
4	Datos complementarios	Δ
5	Posibles análisis para realizar	
6	Diseño conceptual	
7	Diseño lógico relacional	8
8	Implementación y carga	9
8	3.1 Implementación	9
8	3.2 Carga de datos	11
9	Script SQL	22
10	Descarga en formato CSV	
11	Bibliografía	
Tal	bla de ilustraciones	
	stración 1: Petición de información ampliada	
	stración 2: Modelo Conceptual Catastros Turísticos	
	stración 3: Modelo Lógico Relacional Catastros Turísticos stración 4: Quitar valores duplicados	
	stración 5: Reemplazo de vacíos por null	
	stración 6: Modelo Relacional en Workbench	
	stración 7: Exportación del modelo relacional	
	stración 8: Opciones de exportación SQL	
	stración 9: Filtro de exportación de objetos SQL	
	stración 10: Revisión del script generado	
llus	stración 11: Script en Workbench	14
llus	stración 12: Aplicación de restricciones de integridad	15
llus	stración 13: Ejecución del script en MySQL	15
	stración 14: Generación de insert	
	stración 15: Estructura de carga de datos en Visual Studio Code	
	stración 16: Estructura tabla Cantones	
	stración 17: Carga de datos en Workbench	
	stración 18: Creación de usuario con sus permisos	
	stración 19: Generación de tablas y carga de datos correcta	
	stración 20: Visualización de la tabla cantones	
	stración 21: Conexión al usuario proyectoi	
llus	stración 22: Probar conexión al usuario "proyectoi"	20

Introducción

El turismo es una de las ocupaciones más demandadas del sector productivo, ya que

una vez que un turista nacional e internacional llega a un destino, hay varios beneficiarios,

tanto de los sectores como son los hoteles, restaurantes, aerolíneas, que en muchas

ocasiones por falta de publicidad se desconoce que existen, así como de sectores que no

pertenecen al medio turístico como artesanos, transportistas, taxistas, vendedores

ambulantes, etcétera. Los cuales, de más mayor o menor forma, se benefician

económicamente de las ocupaciones que los turistas hacen.

Es por lo que en base a conocer el impacto que tiene el turismo sobre un determinado

territorio el presente proyecto integrador de saberes está orientado a realizar análisis de

datos sobre información turística de los diferentes establecimientos del Ecuador, para lo

cual se aplicó técnicas de exploración y visualización de la información, técnicas de

procesamiento de datos, entre otras, que se adquirieron durante las semanas del curso

académico en el componente de base de datos.

Además, con los resultados obtenidos del procesamiento de datos se pretende obtener

estadísticas que aporten información y conocimiento relevante respecto a la situación en

la que se encuentra la actividad turística en nuestro país.

Finalmente, este documento describe específicamente todo el proceso de

procesamiento y depuración de datos, diseño e implementación de la base de datos

referente al Catastro Turístico de Establecimientos a nivel nacional.

Repositorio GitHub

Url repositorio: https://github.com/Castillo3004/ProyectoIntegradorBDA2022.git

3

3 Datos base

Para este proyecto se nos facilitó un conjunto de datos que tiene disponible el Ministerio de Turismo denominado Catastro Turístico de Establecimientos a Nivel Nacional en todas las actividades turísticas.

El dataset incluye información acerca de todos los establecimientos juntamente con su nombre, dirección, referencias, dirección web, entre otros campos; con los que cuenta el Ecuador y que ofrecen servicios de alimentación, alojamiento, intermediación, transporte, etc., en cada una de las provincias y de cantones de nuestro país.

• Fuente: Ministerio de Turismo

• Conjunto de datos base: https://bit.ly/34luVCd

• Volumen: 22209 registros

Para contar con una data más completa se adjuntó información complementaria que cuenta con variables que nos permitan analizar desde otra perspectiva el catastro. Se detallan a continuación.

4 Datos complementarios

Dentro del conjunto de variables que se adicionaron con la finalidad de implementar una base de datos completa acerca de los establecimientos turísticos de nuestro país tenemos las siguientes:

 Guías turísticos: esta información se le implementó con la finalidad de proporcionar una estadística sobre los guías turísticos con los que cuenta cada provincia de nuestro país. Dentro de esta data obtenida desde el portal (MINISTERIO DE, 2022) se decidió trabajar solo con las siguientes columnas: nombres, apellidos e idioma.

- Ciudades: consideramos a esta variable importante ya que necesitamos obtener un catálogo de las ciudades para conocer con cuantos establecimientos cuenta cada ciudad.
- Latitud: esta variable es muy importante para obtener las coordenadas geográficas de un lugar, dicha variable se la agregó dentro de la entidad ciudades para tener un catálogo con información completa.
- Longitud: también es una coordenada geográfica, que juntamente con la latitud, nos permiten ubicar un punto determinado sobre la superficie terrestre, al igual que la variable latitud esta también pasó a formar parte de la entidad ciudades con el fin de tener un catálogo completo de ciudades.
- Población: consideramos integrar la variable población con la finalidad de obtener estadísticas del total de la población que existe en cada ciudad y comparar la cantidad de establecimientos por cada ciudad.

Para adjuntar las siguientes variables complementarias a la data principal se redactó un correo electrónico a la dirección <u>catastros@turismo.gob.ec</u> con la finalidad de obtener información nueva que nos sea de utilidad para relacionarla con los datos que ya teníamos.



Ilustración 1: Petición de información ampliada

En respuesta a la petición realizada por correo electrónico pudimos dejar finalmente las siguientes variables:

- Estado del registro del establecimiento: esta información es importante y
 decidimos implementarla porque nos permitirá conocer el estado en el que se
 encuentra el centro comercial, es decir conocer si aún se encuentra funcional dicho
 establecimiento.
- **Sistema de origen:** y finalmente esta variable nos permite generar estadísticas sobre el sistema en el que más se encuentran registrados establecimientos comerciales de acuerdo con la provincia donde se encuentre ubicado.

Con esta información pudimos obtener cinco diferentes sistemas de origen que son:

- SIIT: Sistema Integrado de Información Turística
- **SITURIN:** Sistema de Turismo Inteligente
- **SIETE:** Sistema Integrado Estratégico de Turismo del Ecuador.
- **SICET:** Sistema de Catastro de Establecimientos Turísticos

5 Posibles análisis para realizar

- Generar una estadística del número de establecimientos comerciales por provincia que solo pertenezcan al servicio de alojamiento.
- 2. Proyectar estadística de la cantidad de establecimientos comerciales que existen por cada cantón.
- Elaborar estadística de los establecimientos comerciales que pertenecen al sistema de registro SIETE.
- 4. Generar una estadística de las actividades comerciales que más repiten en los establecimientos turísticos.
- Elaborar estadística de los establecimientos que pertenezcan al tipo CENTRO DE TURISMO COMUNITARIO y en qué provincia se encuentran.

- 6. Elaborar estadística de los establecimientos comerciales que se encuentren en la ciudad de Quito que ofrezcan el servicio de parques de atracción estables.
- Proyectar estadística del número de guías turísticas con los que cuenta cada establecimiento.
- 8. Elaborar estadística de los establecimientos comerciales según su categoría.
- 9. Generar estadística de los establecimientos que cuenten con un mayor número de servicios.
- 10. Elaborar estadística de los establecimientos turísticos que pertenecen a la provincia de Loja y no cuentan con una dirección web.

6 Diseño conceptual

El primer paso para empezar a diseñar nuestra base de datos es realizar nuestro modelo conceptual, en cual identificamos 10 entidades de las cuales en su totalidad identificamos relaciones binarias es decir solo relaciones entre dos entidades.

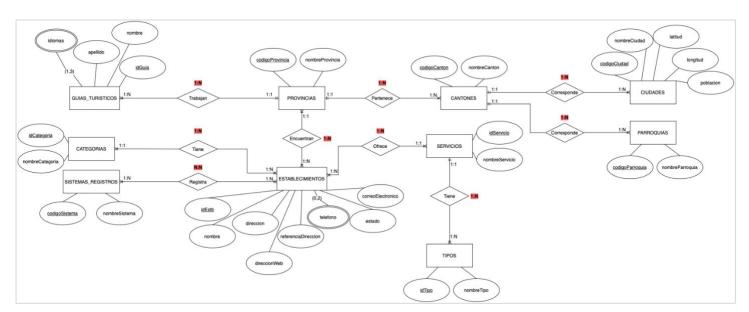


Ilustración 2: Modelo Conceptual Catastros Turísticos

DICCIONARIO DE DATOS

Restricción de dominio adicional

Tipo Entidad /Relación	Atributo	Restricción de dominio adicional	Valor por defecto
Provincias	codigoProvincia	Solo acepta valores numéricos	
Ciudad	codigoCiudad	Solo acepta valores numéricos	
Canton	codigoCanton	Solo acepta valores numéricos	
Guias_turisticos	idioma		Español
Ciudades	poblacion	Debe ser mayor a cero	

7 Diseño lógico relacional

Para el modelo lógico se trabajó en el entorno MySQLWorkbench para posterior facilitar la generación del script.

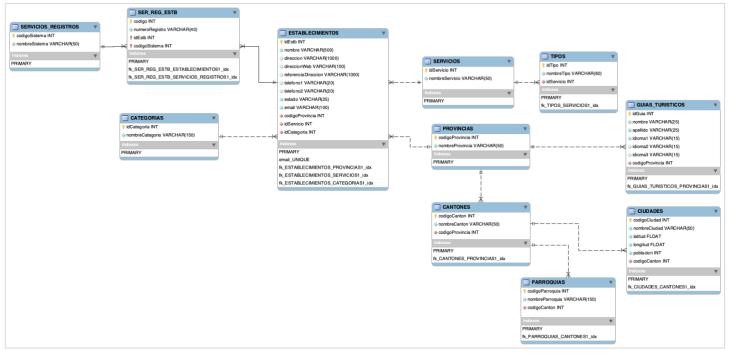


Ilustración 3: Modelo Lógico Relacional Catastros Turísticos

DICCIONARIO DE DATOS

Restricción de dominio adicional

Tabla	Campo	Restricción especifica de dominio	Valor por defecto
Provincias	codigoProvincia	Solo acepta valores numéricos	
Ciudad	codigoCiudad	Solo acepta valores numéricos	
Guias_turisticos	idioma		Español
Ciudades	Población	Debe ser mayor a cero	
Canton	codigoCanton	Solo acepta valores numéricos	

8 Implementación y carga

8.1 Implementación

Para realizar el saneado de datos se utilizó la herramienta Excel por la infinidad de funciones que esta nos ofrece para hacer una depuración eficaz de datos.

- La función VLOOKUP en Excel nos ayudó en gran medida para encontrar valores que se encuentran dentro de un rango de celdas, la búsqueda de esta función se la realiza sobre los valores ubicados en una columna, por lo que también se lo conoce como una búsqueda vertical y es la razón de por qué el nombre de la función inicia con una letra V.
- Para suprimir valores duplicados se aplica en valores del rango de celdas o de la tabla en sí, para otros valores fuera del rango de las celdas o de las tablas no cambian ni se mueven

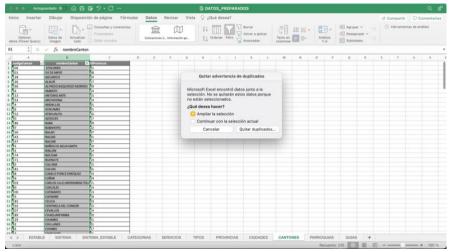


Ilustración 4: Quitar valores duplicados

 Para sustituir un texto se utilizó la herramienta de buscar y reemplazar en Excel para buscar elementos en un libro como números o cadenas de texto en particular.

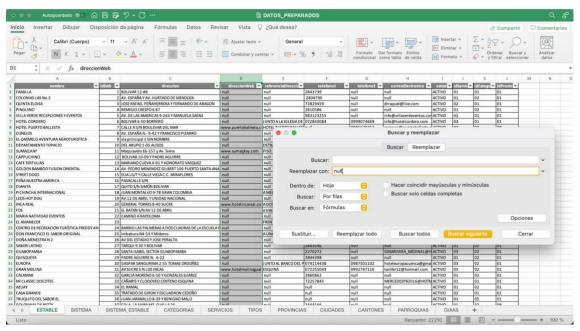


Ilustración 5: Reemplazo de vacíos por null

- Para buscar y reemplazar números o texto se utilizó la herramienta de reemplazar de en Excel.
- Para realizar el filtrado de valores únicos y eliminar los duplicados se realiza un proceso similar ya que el objetivo es presentar una lista de valores únicos, Al momento de realizar el filtrado existe un problema con los valores duplicados y es que al momento de filtrar valores únicos los duplicados se ocultan de manera temporal, sin embargo, al momento de eliminar los valores duplicados estos se eliminan de forma permanente.

8.2 Carga de datos

1. Primero con la herramnienta de MySQL Workbench se procedio a realizar el modelo entidad- relacion, para luego poderlos traducir a lenguaje MySQL, y que nos genere un Script con la creacion de la base de Datos y las tablas con sus respectivas columnas, relaciones, llaves primarias y foraneas, etc.

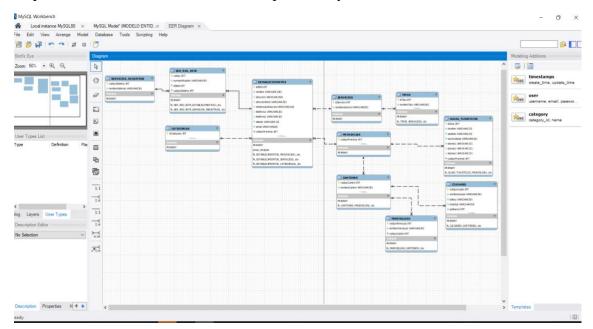


Ilustración 6: Modelo Relacional en Workbench

2. Una vez finalizado el modelo entidad-relacion, nos dirigimos a la pestaña *File*, seguido de esto en la opcion *Export* y aquí podemos seleccionar la opcion de *Forward Engineer SQL CREATE Script*, o si se lo desea hacer de una manera mas rapida, podemos oprimir las teclas Ctrl+Mayus+G.

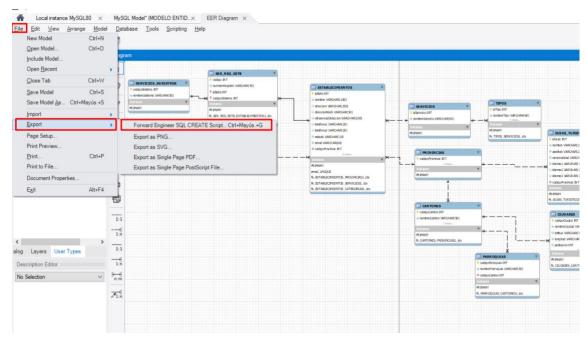


Ilustración 7: Exportación del modelo relacional

3. Entonces, se nos desplegara una ventana con una serie de opcion para el Script SQL, marcamos las dos primeras opciones en este caso, y le damos a continuar.

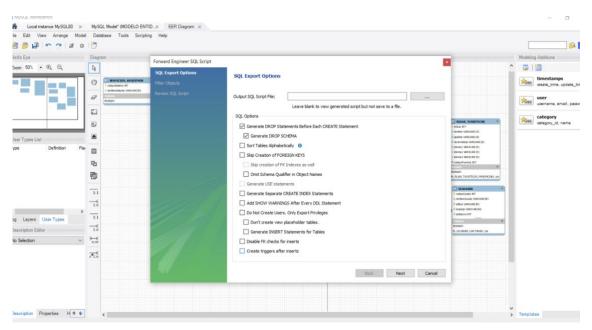


Ilustración 8: Opciones de exportación SQL

4. Aqui verificamos que este marcada la primera opcion y le damos a Continuar.

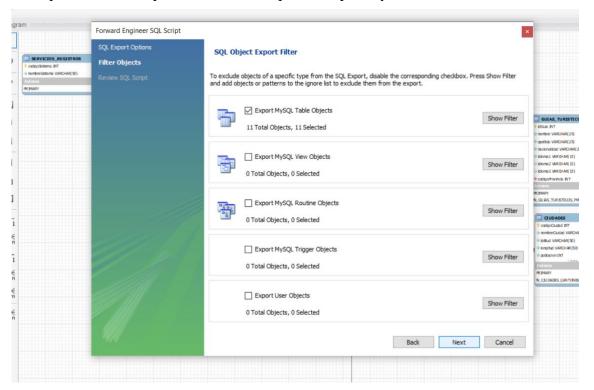


Ilustración 9: Filtro de exportación de objetos SQL

5. Y finalmente nos genera el Script con todo las tablas y sus atributos que habiamos creado en el modelo entidad-relacion, y tambien la creacion de una base de datos con el nombre 'Proyecto', para poder ejecutar este Script, podemos guardarlo o copiarlo.



Ilustración 10: Revisión del script generado

6. En nuestro caso lo guardamos al Script, y procedemos a abrirlo.

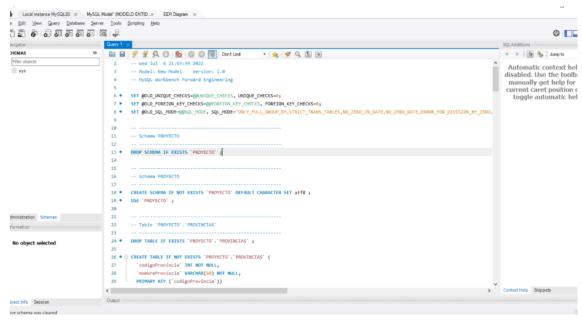


Ilustración 11: Script en Workbench

7. Antes de ejecutar el Script agregamos algunas restricciones, conforme esta establecido en el diccionario de datos.

```
-- Table `PROYECTO`.`GUIAS_TURISTICOS`

    ○ CREATE TABLE IF NOT EXISTS `PROYECTO`.`GUIAS_TURISTICOS` (
    `idGuia` INT NOT NULL,
    `nombre` VARCHAR(25) NOT NULL,
     apellido` VARCHAR(25) NOT NULL,
    `idioma1` VARCHAR(15) NULL DEFAULT 'Español',
    `idioma2` VARCHAR(15) NULL
    `idioma3` VARCHAR(15) NULL,
    `codigoProvincia` INT NOT NULL,
    PRIMARY KEY ('idGuia'),
    INDEX `fk_GUIAS_TURISTICOS_PROVINCIAS1_idx` (`codigoProvincia` ASC) VISIBLE,
    CONSTRAINT `fk_GUIAS_TURISTICOS_PROVINCIAS`
      FOREIGN KEY (`codigoProvincia`)
      REFERENCES `PROYECTO`.`PROVINCIAS` (`codigoProvincia`)
      ON DELETE NO ACTION
      ON UPDATE NO ACTION)
  ENGINE = InnoDB;
```

Ilustración 12: Aplicación de restricciones de integridad

8. Ahora lo ejecutamos a todo el Script y verificamos que nos ha creado la base de datos, junto con sus tablas, columnas, atributos y sus relaciones.

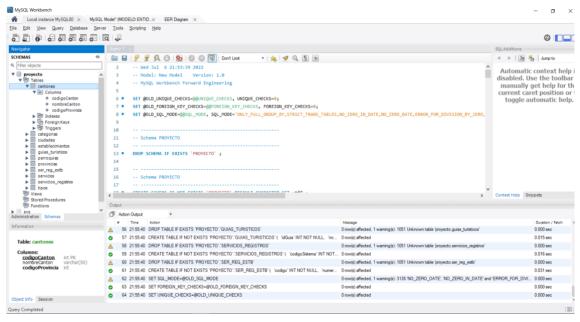


Ilustración 13: Ejecución del script en MySQL

9. Ahora para la carga de datos que antes ya se trabajaron, primero se ha procedido Luego de que los datos han sido saneados correctamente, se ha procedido con Excel a realizar a travez de una funcion que es "&", concatenar las celdas para generar los insert que nos permita realizar la carga de datos a la tabla. ya con la estructura que debe tener que son los parentecis, las comas, y comillas para datos tipo Varchar, esto se realizo para cada una de las tablas.

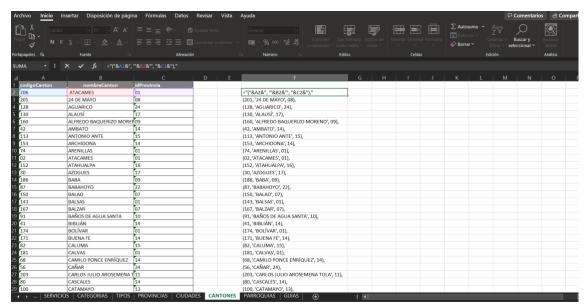
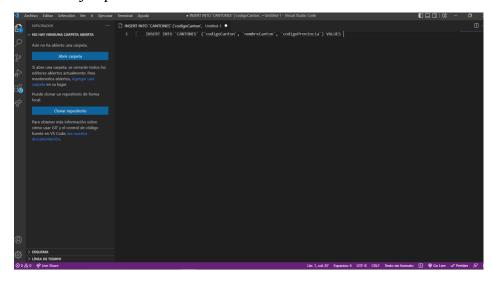


Ilustración 14: Generación de insert

10. Ahora con la ayuda de Visual Studio Code, en un documento de texto en blanco, se escribió la sintaxis de MySQl para la carga de datos en las tablas, en este caso se realiza un ejemplo con la tabla Cantones.



11. Una vez ya con esta estructura, se copian los datos que antes ya se concatenaron.

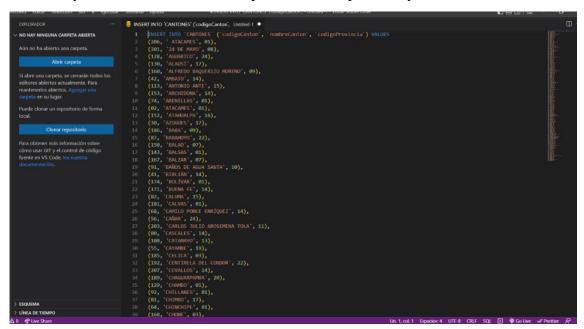


Ilustración 15: Estructura de carga de datos en Visual Studio Code

12. Una vez copiados en el editor de Visual Studio Code, se procede a cambiar la ultima fila de datos por un punto y coma (;), para poderlo pasar todo al Script de MySQL Workbench.

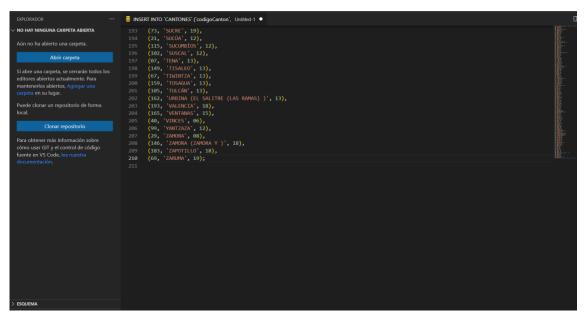


Ilustración 16: Estructura tabla Cantones

13. Ya teniéndolos en MySQL Workbench, se procede a realizar lo mismo con todas las tablas.

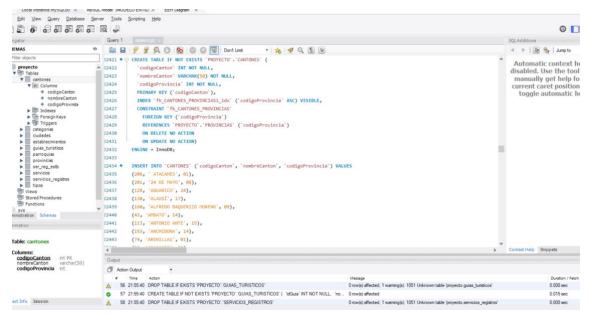


Ilustración 17: Carga de datos en Workbench

14. Luego de esto, se fueron subiendo tabla por tabla los datos, para comprobar que todo se iba subiendo correctamente, y a la final, ya se adjuntaron todos los datos en un solo Script, tanto tablas como datos. Finalmente se procedió a crear un usuario con el nombre de 'proyectoi', contraseña 'mysql' y con los permisos de poder Insertar datos, Borrar Datos, Visualizar los Datos, y poder Subir Datos

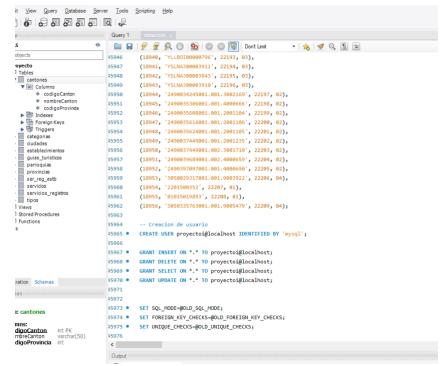


Ilustración 18: Creación de usuario con sus permisos

15. Aquí podemos notar que si se han cargado los datos.

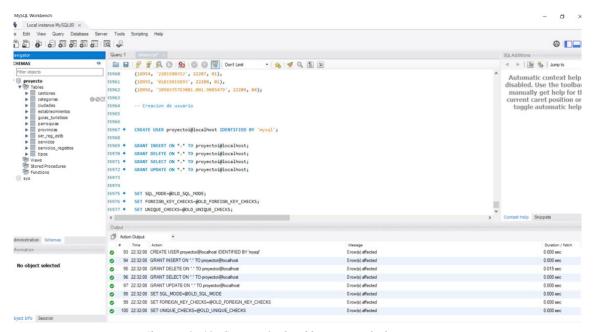


Ilustración 19: Generación de tablas y carga de datos correcta

16. Finalmente se comprobó que los datos efectivamente están subidos, y también el usuario, se lo hizo tanto en MySQL Workbench, y para una mejor seguridad también se lo realizo en IDE de DataGrip.

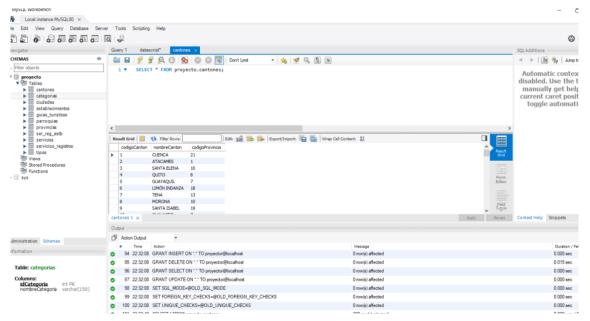


Ilustración 20: Visualización de la tabla cantones

17. Se realización la conexión al usuario que se creó para la gestión de la base de datos, ingresamos tanto el usuario como la contraseña que definimos. Como vemos en la siguiente captura la conexión se realiza de manera exitosa y por ende ya tiene acceso a la base de datos PROYECTO.

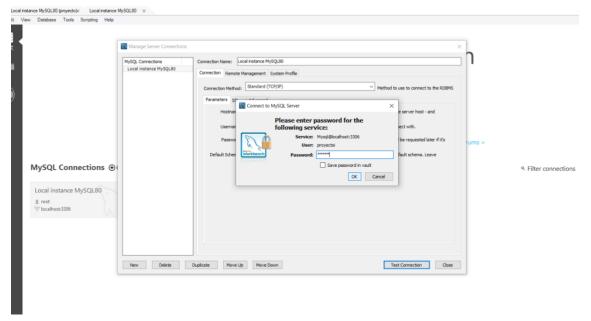


Ilustración 21: Conexión al usuario proyectoi

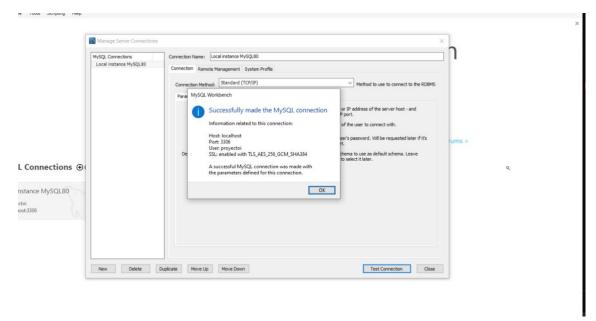
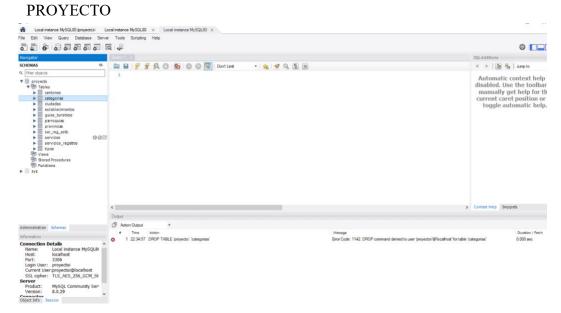
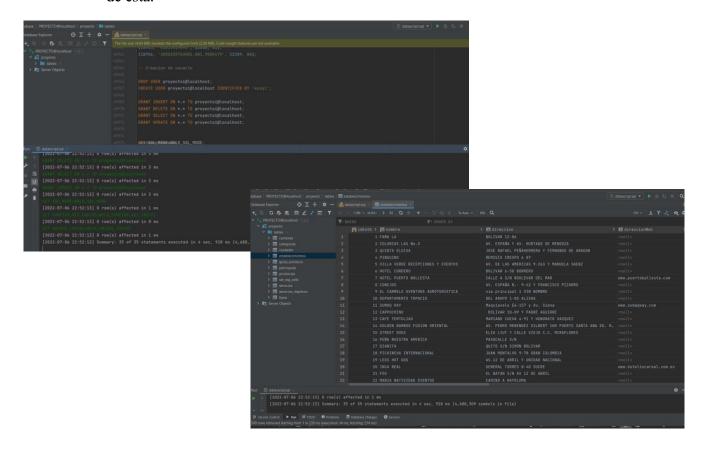


Ilustración 22: Probar conexión al usuario "proyectoi"

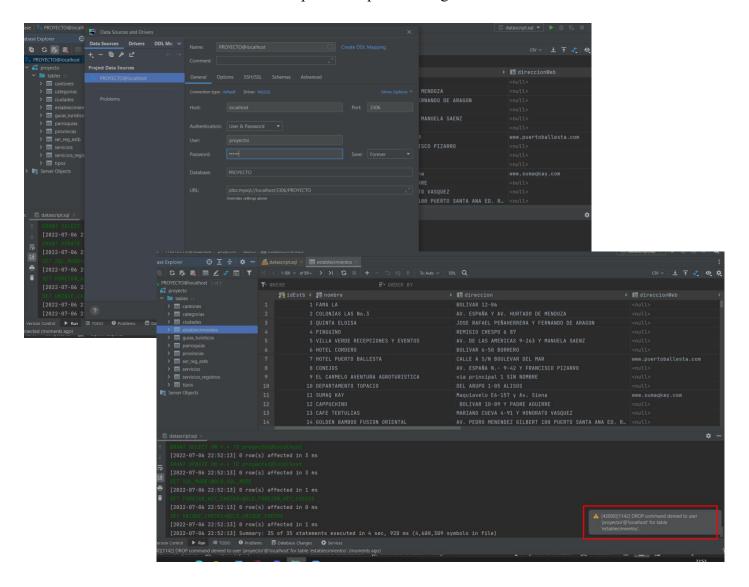
18. Observamos que el usuario que creamos tiene acceso a la base de datos



19. Las capturas que se muestran a continuación son del entorno de DataGrip que es un entorno que se usó para realizar introspección a la base de datos que estábamos creando y obtener mayor seguridad de todo el proceso de creación y carga de datos de esta.



20. Se probo que la conexión al usuario que creamos se realice sin problemas y finalmente se verifico los permisos que se le asigno al usuario.



9 Script SQL

Url Script:

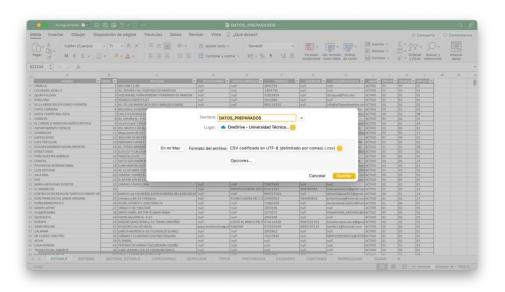
https://github.com/Castillo3004/ProyectoIntegradorBDA2022/blob/181e9b572ecd3099c 29e5ab691a9eb6a7beddad7/Base%20De%20Datos%20Avanzada/datascript.sql

10 Descarga en formato CSV

Haciendo uso de la herramienta Excel, esta nos permitió poder convertir todos los datos procesados que se usaron para la implementación de la base de datos a archivos en formato CSV ya que esta herramienta cuenta con la opción de poder exportar los datos en formato dicho formato.

Una vez que tenemos la data procesada realizamos el siguiente procedimiento que es bastante sencillo para la conversión de cada uno de los catálogos a CSV.

- Archivo -> Guardar como
- En el menú "Formato de archivo", seleccione "CSV codificado en UTF-8
 (delimitado por comas) (.csv)"



URL ARCHIVOS CSV:

https://github.com/Castillo3004/ProyectoIntegradorBDA2022/tree/main/Base%20De%20Datos%20Avanzada/csv

11 Bibliografía

MINISTERIO DE, T. (13 de 06 de 2022). *SERVICIOS TURISMO*. Recuperado el 07 de 2022, de https://servicios.turismo.gob.ec/portfolio/catastro-turistico-nacional