

**UNIVERSIDAD DE SAN CARLOS DE GUATEMALA**  
**FACULTAD DE INGENIERIA**  
**ESCUELA DE CIENCIAS Y SISTEMAS**  
**ESCUELA DE VACACIONES DICIEMBRE 2021**  
**CATEDRATICO: ING. OSCAR PAZ**  
**AUXILIAR: MARIO MORALES**



## **PRÁCTICA 1**

### Objetivos

- Implementar un modelado de datos de acuerdo a los requerimientos dados.
- Realizar la carga masiva de datos en base a los archivos entregados.
- Entender y manipular un modelo relacional diseñado por un tercero.
- Ejecutar sentencias de SQL y scripts mediante archivos.

### Aplicación de Comandos

IDA, la Asociación Internacional de Fomento (por sus siglas en inglés International Development Association) una entidad que trabaja dentro del Banco Mundial que ayuda a los países más pobres del mundo. Supervisado por 173 naciones accionistas, IDA tiene como objetivo reducir la pobreza a través de préstamos (llamados "créditos") y donaciones para programas que impulsen el crecimiento económico, reduzcan las desigualdades y mejoren las condiciones de vida de las personas.

The International Bank for Reconstruction and Development (IBRD), fue establecido por el banco mundial para trabajar como una empresa auto sostenible y proporciona préstamos y asesoramiento a los países de ingreso mediano y los países pobres con capacidad crediticia. IDA y IBRD comparten el mismo personal y las oficinas centrales, para evaluar los proyectos con los mismos estándares rigurosos. La IDA es una de las mayores fuentes de asistencia para los 77 países más pobres del mundo, 39 de los cuales están en África, y es la principal fuente de fondos de los donantes para los servicios sociales básicos en estos países.

Debido a que IDA y IBRD han trabajado con diferentes fuentes de información, se ha decidido que es momento de manejar una sola base de datos que soporte los nuevos retos que las asociaciones enfrentan para trabajar en conjunto. Es por ello que se le ha contratado a usted para desarrollar una solución que integre la información que por el momento se ha trabajado de forma separada y poco eficiente esta información se ha trabajado mediante archivos CSV en los cuales podrá encontrar toda la información necesaria para el abastecimiento de la base de datos.

## Carga Masiva de Datos

IDA maneja distintos archivos para almacenar su información de forma separada en los diferentes departamentos de la asociación:

- **geonameess.csv:** Contiene datos geolocacionales según código geoname.
- **projects.csv:** Contienen toda la información relevante a los proyectos que la asociación tiene y ha tenido.
- **level\_1a.csv:** Contiene información acerca de las rondas de financiamientos que se les han dado a los diferentes proyectos por lugar.

IBRD maneja una pequeña base de datos para almacenar la información relevante a las transacciones de dinero:

- **transactions.csv**

Además, se manejan estándares ISO para los códigos de países y nombres geográficos:

- **country\_codes.tsv:** Contiene información estandarizada sobre países y sus códigos.
- **locations.csv:** Contiene información sobre los tipos de lugares.

Dicha carga masiva debe de contener los siguientes requerimientos y debe de realizarse de la siguiente manera:

1. Se debe de generar una vista de la base de datos en base al modelo relacional del sistema. Esto lo deben de realizar mediante el programa ORACLE DATA MODELER y cada llave primaria de cada entidad debe ser auto incremental es decir cada vez que se inserte un dato, el valor de la llave primaria debe de incrementarse sin necesidad de ingresar el valor de la llave primaria al momento de la inserción.
2. Se debe de crear un Script de Eliminación de las tablas del modelo relacional del sistema.
3. Se debe crear un Script de tablas temporales, en total deben de crearse 6 tablas temporales ya que, en cada una de estas tablas, se realizará la carga masiva de los archivos.csv que se adjuntaron. Una tabla por cada archivo respectivo.

4. Se deben de crear 6 archivos de control, estos archivos son los que, al ejecutarse, realizan la carga masiva de los datos que están en los archivos .csv
5. Un Script que llene las tablas principales que tome los datos que poseen las tablas temporales creadas y cargadas de los puntos anteriores. Este Script posee inserts y querys que según crean necesarios para llenar las tablas principales.

Se debe de tomar en cuenta los siguientes datos totales de cada tabla que deberían de quedar después de llenar las tablas:

- PROJECT: 5881
- TRANSACTION: 11343
- CURRENCY: 1
- TYPE\_STATUS: 2
- LOCATION: 667
- GEONAME: 61243
- LEVEL\_1A: 61243
- COUNTRY\_CODES: 257

Un archivo .bat que ejecute todos los Scripts y archivos de control en el siguiente orden:

- Script de eliminación de tablas.
- Script de creación de base de datos.
- Script de tablas temporales.
- 5 archivos de control.
- Script de llenado de base de datos.

Se Deberá Implementar 2 ROLES uno de Administrador y otro de empleado, el administrador será el encargado de realizar la carga masiva y el empleado únicamente tendrá permiso de realizar consultas. Además, se deben realizar las siguientes consultas

1. ¿Cuántos proyectos iniciaron en el 2011?
2. ¿Cuál fue el proyecto de mayor duración?
3. ¿Cuál fue el proyecto de menor duración?
4. Top 5 de países con más proyectos asignados
5. Top 5 de países con menos proyectos asignados

6. ¿Cuál fue el proyecto con más costo?
7. ¿Cuál fue el proyecto menos costoso?
8. Top 5 de proyectos más costosos
9. Costo Total de Todos los Proyectos
10. Top 3 de las Transacciones más Costosas RI-East Africa Public Health Laboratory Networking Project

## ENTREGABLES

- Todos los archivos utilizados para realizar la carga masiva de los datos.
- Archivos .sql de creación y eliminación de tablas del modelo relacional.
- Archivos de carga de los datos al modelo relacional.
- Screenshots con el total de datos de cada tabla.
- El DBMS a utilizar es Mysql de Linux.
- Modelo Entidad Relación creado en Data Modeler Como Imagen
- URL video YouTube 5 >= video <= 10 minutos

## RESTRICCIONES

- No se permite realizar la carga masiva con ningún asistente gráfico que provea el DBMS.
- La carga masiva debe realizarse mediante el procesamiento de archivo de control o configuración, que contengan la localización de los archivos con los datos a cargar.
- La carga masiva NO DEBE REALIZARSE MEDIANTE EL ASISTENTE DE MYSQL (desde la terminal).
- El sistema operativo deberá ser Linux en la distribución que han venido manejando.

## PENALIZACIONES

- Si se detecta que el código de la base de datos fue generado mediante un dbms y no manual su nota será de 0 puntos.
- Quien no cumpla con las restricciones especificadas no tendrá derecho a la calificación de la práctica.
- Los datos cargados a los modelos deben ser consistentes ya que esto se tomará en cuenta al calificar.
- Copias encontradas, tendrán nota de 0 pts, y serán reportadas a la escuela de sistemas.

**LA ENTREGA DEBE REALIZARSE EL JUEVES 18 DE DICIEMBRE A LAS 23:59 A TRAVÉS DE UEDI. (NO HABRÁ PRORROGA)**