Practica C

Generación Aleatoria de Datos de Prueba

1. Objetivo General

- Conocer y llevar a cabo el proceso para generar datos de prueba de manera aleatoria.
- Aprender el funcionamiento de la aplicación GenerateData.

2. Objetivos Secundarios

- Agilizar la creación de datos de prueba para nuestra base de datos.
- Conocer de forma precisa las funcionalidades y opciones que ofrece GenerateData.

3. Introducción

Cuando llega el momento de probar la implementación de una base de datos, es necesario insertar datos en ella. La tarea de inserción de datos, en muchas ocasiones inventados por el desarrollador, resulta ser una tarea poco productiva, repetitiva y lenta pero a su vez necesaria.

Con el propósito de agilizar y facilitar dicha tarea, se puede recurrir al uso de software especializado en la generación de datos de prueba. Una de estas herramientas es GenerateData. A lo largo del desarrollo de esta práctica se describirá el uso de GenerateData.

3.1. GenerateData

GenerateData es una aplicación que ofrece la creación de datos de forma rápida, dando soporte para generar código, en diferentes lenguajes, para su inserción sin la necesidad de recurrir a un tercero. Esta herramienta es libre y de código abierto, está escrita en JavaScript, PHP y MySQL. GenerateData puede ser usada en línea o de manera local¹.

La principal diferencia entre la versión en línea y la local es que en línea no es posible generar más de 100 datos a la vez, restricción que desaparece en la versión local.

GenerateData ofrece múltiples opciones para generar código, como lo son: LDIF, Excel, SQL CSV, JSON, JavaScript, Perl, PHP, Ruby, XML y HTML.

A continuación se explicará el uso e instalación de esta herramienta.

¹ http://www.GenerateData.com/?lang=es#t2

4. Desarrollo

La intención de esta práctica será utilizar, en primera instancia, la versión en línea, para así conocer el funcionamiento de GenerateData, posteriormente, si la experiencia ha sido de utilidad y atiende las necesidades del usuario, se procederá a describir el proceso de instalación local de la aplicación.

4.1. Ingreso a la página y selección de idioma

El primer paso será acceder a la página de GenerateData:

http://www.generatedata.com/

Esta acción nos llevará a la página de inicio, misma que cuenta con cuatro pestañas: *Generate, About News y Donate*.

Antes de comenzar con el proceso de generación de datos, si nos es conveniente, es posible elegir el idioma en que se desplegará la interfaz de usuario, ver Figura A.1.



Figura A.1 Selección de idioma.

Hecho esto, los campos requeridos para la creación de datos se despliegan, ver Figura A.2.

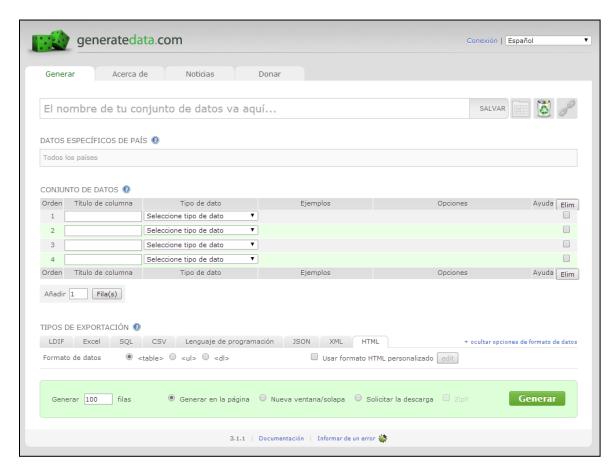


Figura A.2 Página de inicio de GenerateData.

Para esta primera parte de la práctica se trabajará en la pestaña Generate o Generar.

4.2. Nombrado y regionalización del conjunto de datos

A continuación se debe especificar el nombre del conjunto de datos a generar, esto se realiza en la ventana "El nombre del conjunto de datos va aquí...", en este caso lo llamaremos *Prueba*, ver Figura A.3.

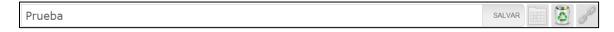


Figura A.3 Nombre del conjunto de datos.

Después, en la sección *Datos Específicos del país* es posible seleccionar el país para "regionalizar" los tipos de datos, el programa contempla ciertos tipos de zonas y países para la generación de datos, ver Figura A.4, para fines de esta práctica se puede elegir cualquiera.



Figura A.4 Selección del país.

4.3. Definición del conjunto de datos

Para generar los datos se tendrá que tomar en cuenta la nomenclatura de los mismos. Ésta consiste en el nombre de la columna y el tipo de dato de acuerdo a lo establecido en el esquema de la base de datos sobre la que se trabaja, ver Figura A.5.

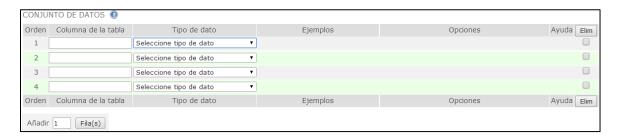


Figura A.5 Conjunto de datos.

Los campos a considerar son los siguientes:

a. Nombre Columna de tabla:

Se coloca el nombre de la columna tal y como aparece en el esquema de la base de datos. Se pueden especificar varias columnas, por default hay opción para cuatro pero si se necesitan agregar más, bastará con dar clic en el botón de Añadir Fila(s), pudiendo añadir las que el usuario desee.

b. Tipo de dato

GenerateData agrupa los tipos de datos en 7 categorías que son: Datos humanos, Geográficos, de Texto, Numéricos, Tarjeta de Crédito, Matemáticos y Otros. Dependiendo de la categoría, a su vez, éstas se descomponen de la siguiente manera:

- i. **Datos humanos** incluye formatos como nombres, teléfono/fax, correo electrónico, fecha, empresa o números de cuenta bancaria.
- ii. **Datos geográficos** contiene formatos para dirección, ciudad, código postal, región, país y latitud/longitud.
- iii. **Datos de tarjeta de crédito** permite definir elementos como el PAN, PIN, CVV, Track1 y Track2.

- iv. **Datos de texto** permite fijar palabras o poner un número aleatorio de éstas.
- v. **Datos numéricos** que pueden ser alfanuméricos, autoincremento, rango numérico, GUID y de concurrencia.
- vi. Datos matemáticos define valores numéricos con distribución normal.
- vii. **Otros** incluye formatos como constante, compuesto, de árbol o lista personalizada para generar datos con un formato específico que se requiera.

c. Ejemplos

Dependiendo del "Tipo de dato" elegido, esta opción mostrara ciertas especificaciones particulares para la nomenclatura de los datos de la columna.

d. Opciones

Esta sección también dependerá del "Tipo de dato" y del formato de "Ejemplos", permitiendo definir el formato, rango de longitudes, fechas y números.

e. Ayuda

La opción de "Ayuda" permite documentarse más sobre el "Tipo de dato" que se seleccionó, describiendo a detalle los datos a generar, ver Figura A.6.

f. Eliminación

Permite eliminar la columna del conjunto de datos.

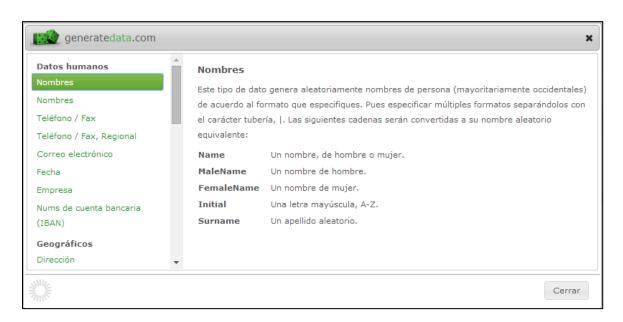


Figura A.6 Ayuda de conjunto de datos.

A manera de ejemplo, a continuación se muestra cómo se generarían datos de prueba para una tabla con cuatro columnas: identificador, nombre, fecha y lugar de nacimiento, ver Figura A.7.

Supongamos que las características de estos cuatro atributos son las siguientes:

• identificador: dato numérico que se incrementa automáticamente de uno en uno, comenzando en 1.

- **nombre:** cadena de caracteres que representa el nombre de una persona sin importar su sexo, nótese que la herramienta permite elegir entre nombres femeninos, masculinos o ambos.
- **fecha_de_nacimiento:** fecha que indica el día, mes y año en que nació una persona, es posible indicar los límites inferior y/o superior de las fechas, para este caso el límite inferior será el año 1900 y el superior la fecha actual.
- pais_de_nacimiento: cadena de caracteres que indicará el país de nacimiento del individuo.

La combinación anterior, reflejada en GenerateData, se puede observar en la Figura A.7.

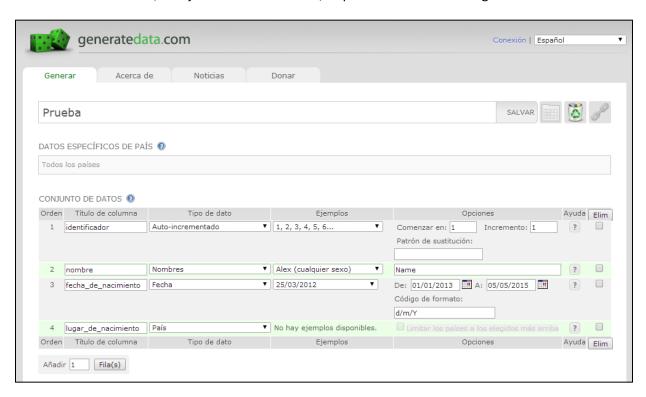


Figura A.7 Generación de Tabla con cuatro columnas.

4.4. Exportación del conjunto de datos

En la sección de "Tipo de exportación" se muestra una serie de opciones para generar el código, ver Figura A.8.



Figura A.8 Tipo de exportación.

GenerateData permite exportar a 8 tipos diferentes, los cuales se describen a continuación:

 a. CSV
Este formato permite generar los datos separados por comas para su manejo y exportación a programas que sean compatibles con este tipo de extensión. Dando la opción para definir el carácter delimitador entre los datos y al carácter de fin de línea compatible para Windows, Linux y Mac.

b. Excel

Genera un archivo para abrir directamente con Excel, a manera de las columnas requeridas y renglones.

c. HTML

HTML es otra forma de dar formato al contenido de las páginas web.

d. JSON

Este formato de salida es de uso común para el intercambio de datos, ligero, fácil de leerlo y escribirlo para las personas y para las computadoras es de fácil manejo². Por lo que GenerateData da dos opciones para la estructura de datos, "Complejo" o "Simple". Permitiendo "Eliminar espacios en blanco en los resultados generados".

e. LDIF

Es un tipo de formato de archivo que sirve de intercambio de datos para transferir la información entre servidores LDAP que permite buscar información en la red por medio del acceso a la información del servicio de directorio de manera legible³.

f. Lenguaje de programación

Esta opción permite guardar el conjunto de datos con formato de JavaScript, Perl, PHP y Ruby.

g. SQL

Para este tipo de archivo, GenerateData ofrece varias opciones, como puede ser el nombre de la tabla para el conjunto de datos, el "Tipo de base de datos" y que son compatibles con MySQL, Postgres, SQLite, Oracle y SQL Server, diferentes opciones de creación de código para la creación de la tabla, "Tipo de sentencia" para el conjunto de datos y "Clave primaria" de autoincremento en los datos, ver Figura A.9.

h. XML

Esta opción permite construir un documento XML de los datos generados.

Para fines de esta práctica se elegirá la pestaña de SQL, ver Figura A.9.



Figura A.9 Tipo de exportación SQL.

³ http://web.mit.edu/rhel-doc/4/RH-DOCS/rhel-rg-es-4/ch-ldap.html

² http://www.json.org/json-es.html

En esta pestaña se tienen opciones como: Nombre de tabla, Tipo de base de datos, Opciones Varias, Tipo de sentencia y Clave primaria, mismas que se detallan a continuación:

- i. **Nombre de tabla:** aquí se especifica el nombre de la tabla de la base de datos en la que se pretenden insertar los datos.
- ii. **Tipo de base de datos:** ofrece diferentes opciones de creación de código como lo es MySQL, Postgres, SQLite, Oracle y SQL Server.
- iii. **Opciones varias:** en primera instancia se da la opción para incluir en el código la creación de la tabla, es decir la sentencia CREATE TABLE. De igual manera otorga la opción para la sentencia DROP TABLE, que en muchos casos es conveniente a la hora de estar probando un conjunto de datos en una base de datos, esta opción borra la tabla si existe y en seguida comienza con las demás sentencias del código y en último lugar se tiene la opción para entrecomillar cada inserción de tabla y columnas.
- iv. **Tipo de sentencia:** la opción INSERT permite insertar nuevos registros en la tabla existente⁴, INSERT IGNORE permite insertar datos, si estos ya han sido creados, el comando IGNORE ignora esos datos e inserta sin restricción, UPDATE actualiza las columnas de la tabla existente en los registros con nuevos valores⁵.
- v. Clave Primaria: la opción NINGUNO dejara sin clave primaria a la tabla, por el contrario la segunda opción añadirá la llave primaria como un atributo numérico de autoincremento.

En este momento es posible salvar el conjunto de datos que se ha definido, esto se hace dando clic en el botón "SALVAR" al lado derecho de donde se puso el nombre del conjunto de datos, ver Figura A.3.

Además de la opción salvar, GenerateData permite tres funcionalidades más, ver Figura A.10:

- 1. Tus conjuntos de datos (ícono de la carpeta cuadriculada): muestra los conjuntos de datos ya creados.
- 2. Limpiar página (cesto de basura): limpia la página, dejando los campos en blanco para una nueva creación de conjuntos de datos.
- 3. Enlazar a este conjunto de datos (eslabones de una cadena): Comparte el conjunto de datos con alguien más.

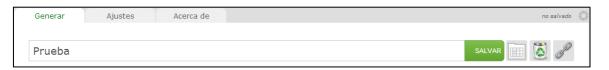


Figura A.10 Opciones de Salvar el conjunto de datos.

Para comprobar que se ha guardado el conjunto de datos, se deberá dar clic en el botón "Tus conjuntos de datos", esta acción desplegará el listado de los conjunto de datos existentes, ver Figura A.11.

⁴ http://dev.mysql.com/doc/refman/5.0/es/insert.html

⁵ http://dev.mysql.com/doc/refman/5.0/es/update.html

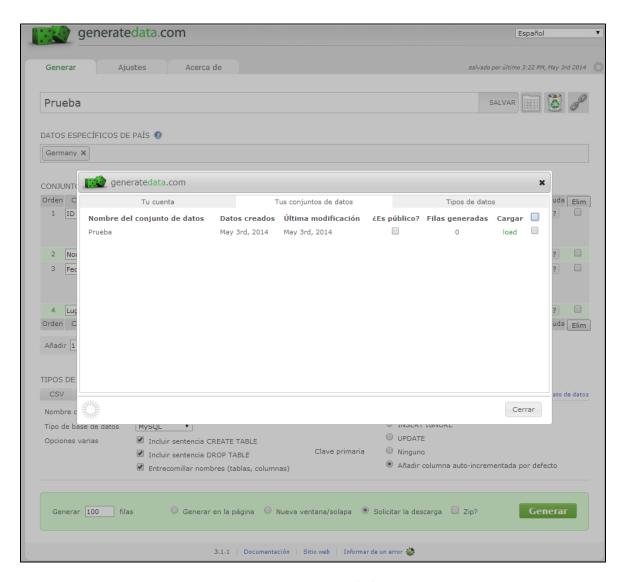


Figura A.11 Tus conjuntos de datos.

Por último se debe especificar la forma de ver los datos, ya sea en el navegador en la misma página, en una ventana nueva o descargando el archivo, ya sea normal o comprimido, finalizando al dar clic en "Generar", ver Figura A.12.



Figura A.12 Generar datos.

A continuación, a manera de ejemplo, se muestra la generación de un conjunto de datos con seis columnas de diferentes tipos de registros, así como la generación de código en MySQL y la opción seleccionada para que los datos se generen en una nueva ventana del navegador, ver Figura A.13 y Figura A.14.

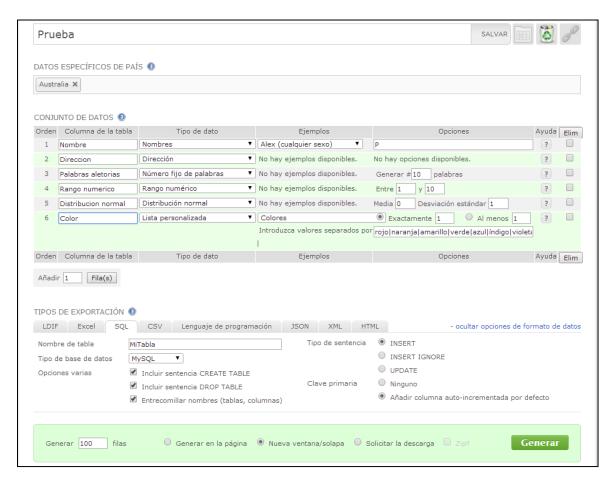


Figura A.13 Creación de registros en GenerateData.

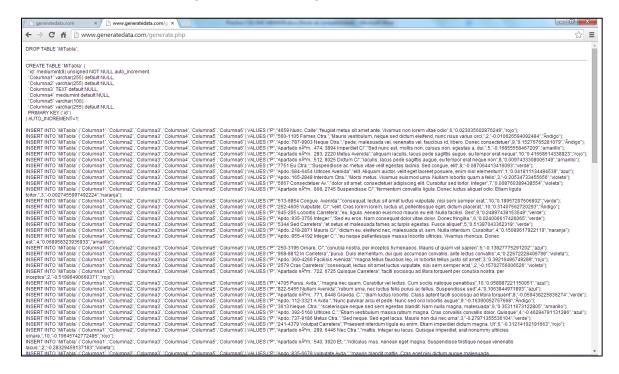


Figura A.14 Generación de código y registros en MySQL.

Finalmente, para poblar la base de datos con el conjunto de datos generado, bastará con copiar el código generado por GenerateData y pegarlo en la consola del SMBD que se esté utilizando.

4.5. Descarga e instalación de manera local

Para la instalación de manera local es necesario cambiar a la pestaña *About* o *Acerca de* y dar clic en el botón Descargar, ubicado en la parte inferior de la página, ver Figura A.15.

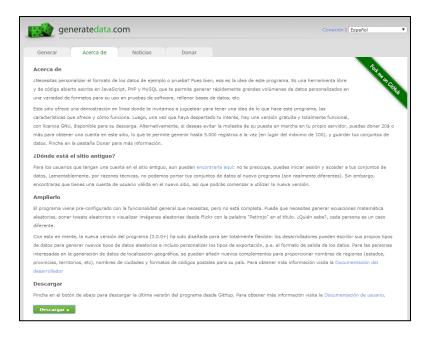


Figura A.15 Descarga de archivo generateData desde la página oficial.

Esto desplegará un listado con las diferentes versiones de GenerateData, para los fines de esta práctica se elegirá la más reciente, en este caso, la versión 3.1.1, ver Figura A.16.

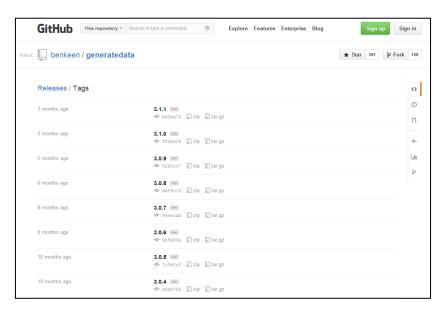


Figura A.16 Descarga de generateData-versión-3.1.1.

Para comenzar la instalación es necesario arrancar los servicios de Apache, MySQL y Tomcat desde XAMPP, ver Figura A.17, con esto se iniciara el proceso para instalar una base de datos, agregar un usuario y darle privilegios.

En caso de requerir mayor información sobre el uso e instalación de XAMPP, consulta la práctica B de este manual.

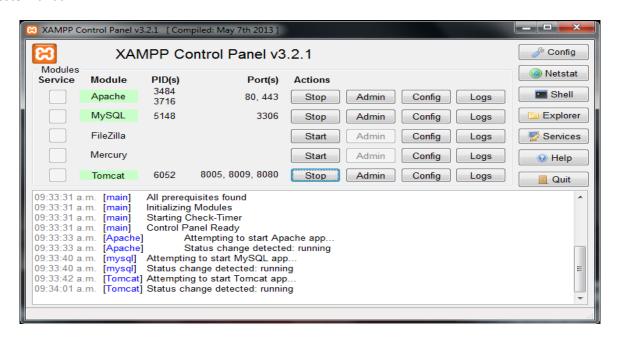


Figura A.17 Inicio de servicios en XAMPP.

Desde el panel de control de XAMPP nos dirigimos a la ventana de administración de MySQL, esto se hace dando clic en el botón "Admin" del servicio MySQL, el cual nos direccionará a dicha ventana en el navegador predeterminado en la máquina, ver Figura A.18.

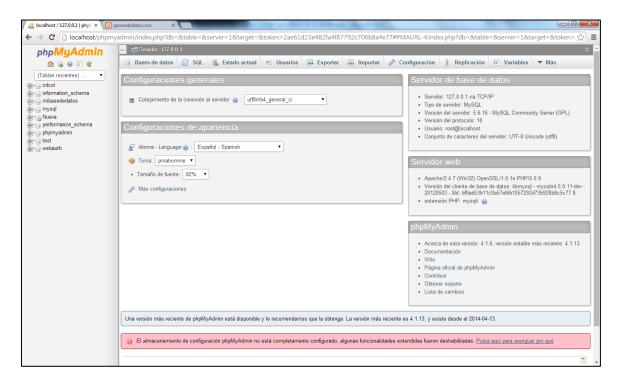


Figura A.18 Administración de MySQL.

Para crear la base de datos, dar clic sobre el botón "Nueva", esto en la parte izquierda de administración, ver Figura A.19.

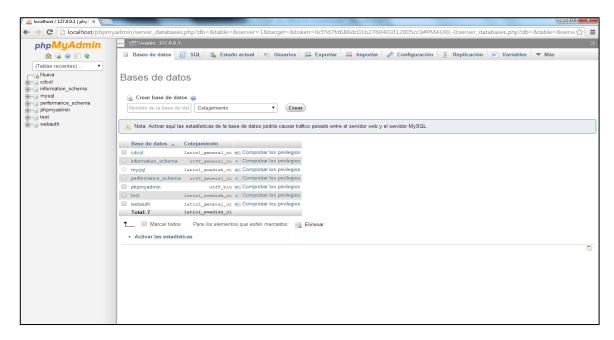


Figura A.19 Creación de base de datos.

Colocamos el nombre de la base de datos en el campo requerido, en este caso "mibasededatos" y clic en el botón "Crear", ver Figura A.20.

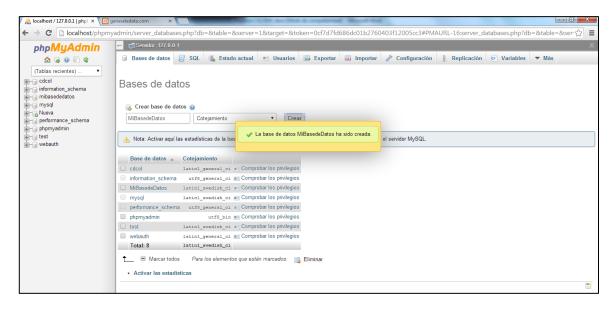


Figura A.20 Creación de la base "mibasededatos".

Seleccionar la base de datos creada, después clic sobre privilegios en la parte superior derecha, ver Figura A.21 y Figura A.22 y clic en "Agregar usuario".

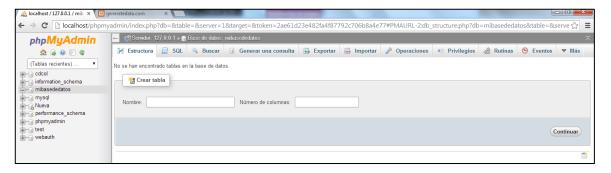


Figura A.21 Selección de la base de datos.

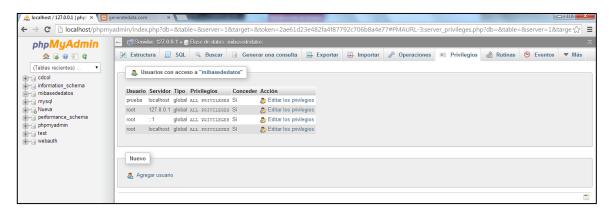


Figura A.22 Agregando usuario.

Dentro de la ventana de Agregar usuario, se deberán llenar los campos requeridos, ver Figura A.23 y Figura A.24.

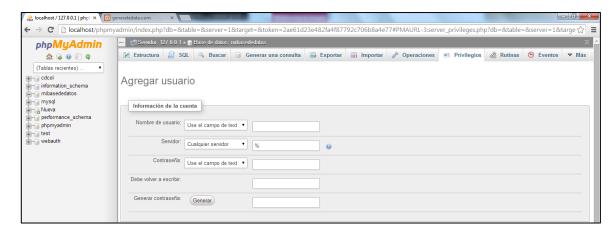


Figura A.23 Requerimientos para el nuevo usuario de la base de datos I.

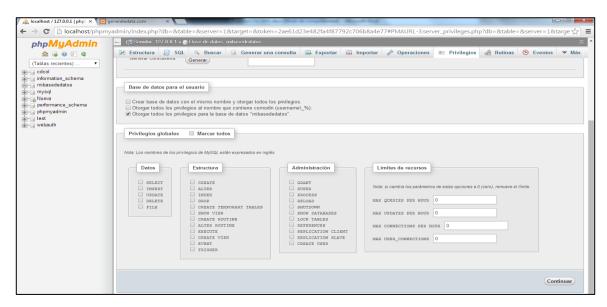


Figura A.24 Requerimientos para el nuevo usuario de la base de datos II.

Para los fines de ésta practica, colocamos en "Nombre de usuario" como usuario, en "Servidor" seleccionamos una conexión local, automáticamente el campo se colocara como localhost, damos una contraseña y la repetimos, en el campo "Generar contraseña" se deja vacío, esta opción es opcional, ver Figura A.25.

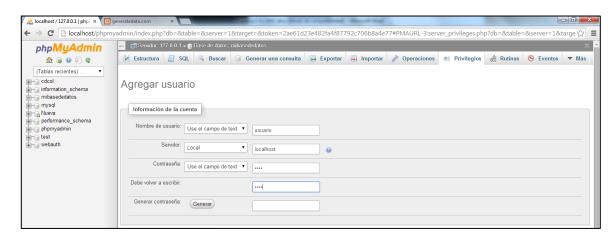


Figura A.25 Campos requeridos para el usuario.

Desplegando la ventana de la creación del usuario hacia abajo, dejamos las opciones que vienen por default en el apartado de "Base de datos para el usuario" y damos permisos totales de:

- Datos: como lo son Selección, Inserción, Actualización, etc.
- Estructura: que serán de Creación, Modificación, Índices, Borrado, Vistas, Eventos, etc.
- **Administración:** para la creación de usuarios, bloqueo de tablas para transacciones⁶, referencias, vista de base de datos, etc.

En los campos de "Limites de recursos" se dejan las opciones por default y se da clic en el botón "Continuar", ver Figura A.26.

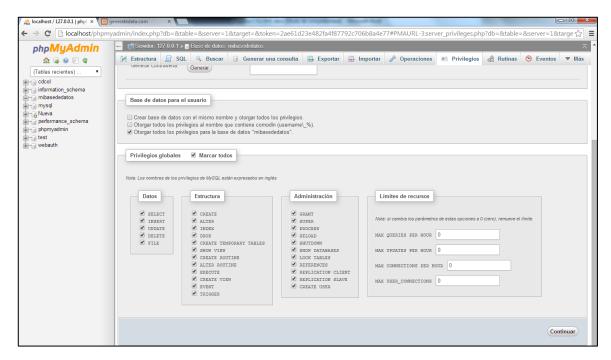


Figura A.26 Privilegios para el usuario.

⁶ http://es.tldp.org/Postgresql-es/web/navegable/user/sql-lock.html

Una vez hecho los pasos se agregará con éxito el usuario, ver Figura A.27.



Figura A.27 Usuario agregado en la base de datos.

Para iniciar GenerateData, se deberá descomprimir el archivo descargado en el segundo paso de esta subsección, ver Figura A.16 y colocarlo en la carpeta "htdocs" de XAMPP, si la instalación de XAMPP fue hecha en la unidad C, entonces la ruta seria C:/xampp/htdocs.

Hecho lo anterior, procederemos a escribir la siguiente dirección en el navegador

localhost/GenerateData-3.1.1/

Con esta acción comenzarán los pasos para la instalación desde el navegador, ver Figura A.28.



Figura A.28 Instalación de GenerateData.

Al llenar los campos se pedirá:

- **Host Name:** en el que se instalará GenerateData, por default coloca la máquina local, "localhost".
- **Database Name:** nombre de la base de datos creada en MySQL, "mibasededatos", ver Figura A.20.
- MySQL Username: el nombre de un usuario que tenga acceso a ella, en este caso "usuario".
- **MySQL Password:** la contraseña del usuario colocado en el apartado anterior, para fines prácticos se colocó "1234".

- **Table Prefix:** es el prefijo que se colocará a todas las tablas necesarias para la instalación de GenerateData. Se deberá dejar la opción por default "gd_". Nótese que estas tablas las generará automáticamente el instalador en pasos posteriores.
- **Default Language:** indica el idioma del instalador, se elige "Español" y se da clic en "Continue", ver Figura A.29.

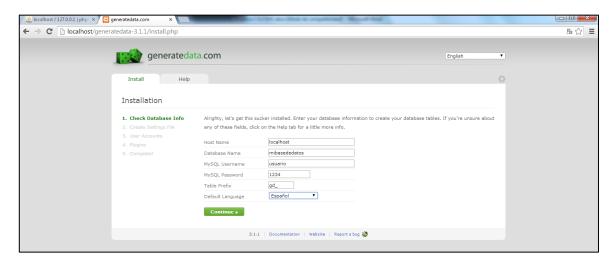


Figura A.29 Llenado de campos en el formulario de instalación.

El siguiente paso será la creación del archivo de configuración "settings.php" para GenerateData, esto se hace dando clic en "Create File", ver Figura A.30.



Figura A.30 Creación del archivo "settings.php".

Después el instalador pedirá seleccionar un tipo de cuenta de instalación⁷, ver Figura A.31:

- Una solo cuenta anónima: esta opción guardará todas las configuraciones en una sola cuenta.
- ii. **Una sola cuenta con inicio de sesión:** esta permite hacer lo mismo que la anterior, con el agregado de que se solicitará el inicio de sesión al entrar a GenerateData.

⁷ http://benkeen.github.io/GenerateData/install.html

iii. **Cuentas múltiples:** permite configurar cualquier tipo de cuentas de usuario, cada uno con su propia información de inicio de sesión.



Figura A.31 Tipo de cuenta para GenerateData.

La instalación de GenerateData ha sido de manera local y lo recomendable será seleccionar la primer opción, "Single, anonymous user account" o "Una sola cuenta anónima" y dar clic en el botón "Continue", ver Figura A.32.

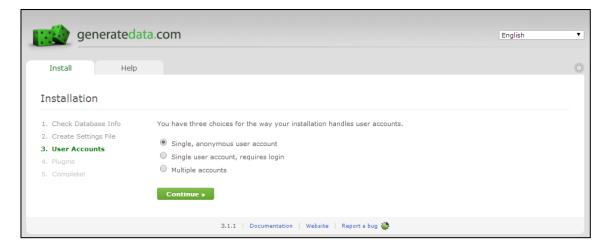


Figura A.32 Creación de la cuenta

A continuación se requerirá instalar los plugins necesarios para generar conjunto de datos, dando clic en "Install Plugins", ver Figura A.33.



Figura A.33 Instalación de plugins.

Comenzará el proceso de instalación de los plugins, una vez termine, dar clic en "Continue", ver Figura A.34. Por medio de un código de colores se indicará si la instalación fue exitosa, verde, o no, rojo.

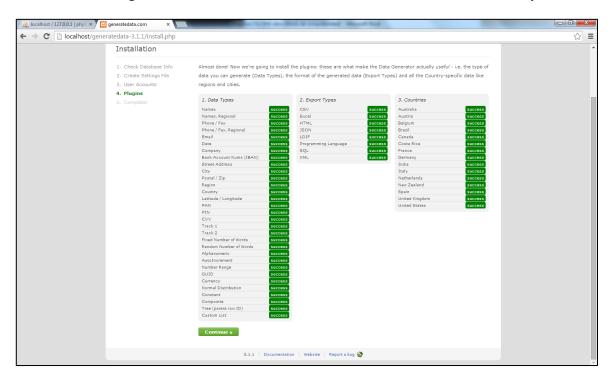


Figura A.34 Instalación exitosa de los plugins

Para terminar dar clic en "Go to script", esto direccionara a la página principal de GenerateData para comenzar con el proceso de generar el conjunto de datos, ver Figura A.35.

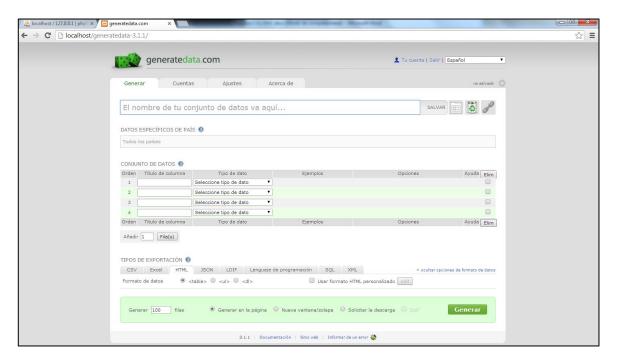


Figura A.35 Página principal para generar conjunto de datos.

5. Ejercicios

- a) Realiza la generación de los conjuntos de datos para poblar tu base de datos con las diferentes opciones de "Conjunto de datos", "Tipo de exportación" y "Generación de datos".
- b) Investiga que otras plataformas ofrecen este tipo de generación de datos de manera automática.
- c) Instala GenerateData en tu computadora. (Opcional)

Entregables requeridos para prácticas subsecuentes:

• Conjuntos de datos para cada una de las tablas de tu base de datos.