

Universidad San Carlos de Guatemala  
Facultad de Ingeniería  
Escuela de Ciencias y Sistemas  
Base de Datos 1  
Ing. Luis Espino  
Aux. Gerardo Chay  
Primer Semestre 2022



## PROYECTO NO. 1

### INTRODUCCIÓN

Muchas organizaciones al carecer de un sistema de información optan por cambiar la manera habitual de guardar sus datos de manera física a una forma virtual. Para estas situaciones es necesario que se diseñe una estructura que guarde la información de forma persistente y de fácil acceso. Uno de los casos que regularmente ocurre es que las empresas pequeñas deciden llevar el manejo de información por medio de archivos de excel, sin embargo, después de un tiempo al aumentar el volumen de datos el manejo de estos archivos se vuelve complicado y muy laborioso, debido a que cada vez el archivo aumenta su tamaño y complejidad.

### OBJETIVOS

- **Objetivo General**

- Que el estudiante sea capaz de diseñar una base de datos a partir de los diferentes modelos por los que se necesita que pase una base de datos.

- **Objetivos Específicos**

- Que el estudiante aprenda a crear un modelo conceptual para una base de datos.
- Que el estudiante aprenda a mapear los elementos de un modelo conceptual y transformarlo en un modelo lógico.
- Que el estudiante pueda crear el modelo físico a partir del modelo lógico que se ha propuesto.
- Que el estudiante utilice el lenguaje SQL para crear reportes útiles para el cliente.

## DESCRIPCIÓN GENERAL

Big Smoke Institution (BSI) es una institución dedicada a registrar, controlar y ayudar a diferentes personas a que dejen su adicción al tabaco. Para esta institución es urgente contar con un sistema de bases de datos donde se puedan hacer consultas de diferentes casos y tratamientos que han sido efectuados a distintos pacientes.

Para la BSI lo más importante son sus empleados y sus pacientes. Para llevar un mejor control de toda la información es necesario que se guarden ciertos datos de estas personas. De los empleados se necesita saber su nombre, apellidos, dirección, teléfono, fecha de nacimiento y su género. De los pacientes se necesita saber lo mismo que un empleado únicamente agregando datos como la altura y el peso que son importantes para algunos análisis. Los empleados dentro de la BSI son profesionales en algún área determinada, es por ello que para la institución es importante saber el título académico que posee cada empleado.

A cada paciente que llega a la BSI se le realiza una evaluación general llevada a cabo por algún empleado. Para el BSI es importante saber en qué fecha se realizó la evaluación. En esa evaluación se determinan todos aquellos síntomas que puede llegar a presentar un paciente. (Ej.: presión alta, ansiedad, depresión, somnolencia, etc.). La BSI maneja un diagnóstico general por cada síntoma que han descubierto en pacientes. Un síntoma puede ser asociado a múltiples diagnósticos y viceversa. El diagnóstico que es asignado a un síntoma es determinado por un rango de 1 a 10, donde 1 representa no tan acertado y 10 como muy acertado.

Una vez finalizada la detección de síntomas y que a cada síntoma se le asigna su diagnóstico basado en un rango, se le indica al paciente que se suministre múltiples tratamientos para superar poco a poco su adicción. El BSI cuenta con diferentes tratamientos categorizados que ya han sido aplicados a otros pacientes y han tenido un impacto positivo. (Ej.: hacer dieta, ejercicio, correr, etc.). Para el BSI es importante saber la fecha en la que su paciente comenzó los tratamientos que le fueron recomendados.

## SOLICITUD

La BSI le proporciona a usted un archivo CSV: [BIG SMOKE DATA](#), que contiene toda la información que necesita almacenar en la base de datos. Una vez analizados los requerimientos (tanto del enunciado como del archivo de carga), se necesita que usted diseñe un nuevo modelo de datos para almacenar toda esa información. El proceso es el siguiente:

1. Debe crear una tabla temporal en donde debe cargar toda la información del CSV. La herramienta que debe utilizar es SQL\*Loader, la cual es propia de Oracle. Su funcionamiento será enseñado en el laboratorio.
2. Haciendo uso de consultas a la tabla temporal debe de llenar el modelo que usted propone para almacenar la información de forma correcta de modo que no exista redundancia de los datos.
3. Una vez lleno todo el modelo, debe realizar los reportes de la sección **REPORTERÍA** los cuales serán evaluados para verificar que la información mostrada sea consistente y verídica.

## REPORTERÍA

1. Mostrar el nombre, apellido y teléfono de todos los empleados y la cantidad de pacientes atendidos por cada empleado ordenados de mayor a menor.
2. Mostrar el nombre, apellido, dirección y título de todos los empleados de sexo masculino que atendieron a más de 3 pacientes en el año 2016.
3. Mostrar el nombre y apellido de todos los pacientes que se están aplicando el tratamiento "Tabaco en polvo" y que tuvieron el síntoma "Dolor de cabeza".
4. Top 5 de pacientes que más tratamientos se han aplicado del tratamiento "Antidepresivos". Mostrar nombre, apellido y la cantidad de tratamientos.
5. Mostrar el nombre, apellido y dirección de todos los pacientes que se hayan aplicado más de 3 tratamientos y no hayan sido atendidos por un empleado. Debe mostrar la cantidad de tratamientos que se aplicó el paciente. Ordenar los resultados de mayor a menor utilizando la cantidad de tratamientos.
6. Mostrar el nombre del diagnóstico y la cantidad de síntomas a los que ha sido asignado donde el rango ha sido de 9. Ordene sus resultados de mayor a menor en base a la cantidad de síntomas.
7. Mostrar el nombre, apellido y dirección de todos los pacientes que presentaron un síntoma que al que le fue asignado un diagnóstico con un rango mayor a 5. Debe mostrar los resultados en orden alfabético tomando en cuenta el nombre y apellido del paciente.
8. Mostrar el nombre, apellido y fecha de nacimiento de todos los empleados de sexo femenino cuya dirección es "1475 Dryden Crossing" y hayan atendido por lo menos a 2 pacientes. Mostrar la cantidad de pacientes atendidos por el empleado y ordénelos de mayor a menor.
9. Mostrar el porcentaje de pacientes que ha atendido cada empleado a partir del año 2017 y mostrarlos de mayor a menor en base al porcentaje calculado.
10. Mostrar el porcentaje del título de empleado más común de la siguiente manera: nombre del título, porcentaje de empleados que tienen ese título. Debe ordenar los resultados en base al porcentaje de mayor a menor.
11. Mostrar el año y mes (de la fecha de evaluación) junto con el nombre y apellido de los pacientes que más tratamientos se han aplicado y los que

menos. (Todo en una sola consulta). **Nota:** debe tomar como cantidad mínima 1 tratamiento.

## RESTRICCIONES

- **DISEÑO:**

- Se debe hacer uso de DataModeler para generar los diseños del modelo relacional.
- Cada modelo debe ser diseñado de manera que evite redundancia de datos y mantenga la integridad de los mismos.

- **SCRIPTS:**

- Para los scripts de las bases de datos el DBMS a utilizar debe ser **Oracle** excepto el script de creación ya que este debe ser creado POR EL ESTUDIANTE, por lo que no puede utilizar ninguna herramienta.
- No se debe separar en diferentes archivos CSV el archivo proporcionado. Debe ser cargado únicamente el que fue proporcionado.
- Los reportes deben ser realizados al modelo relacional creado por el estudiante, reportes hechos a la tabla temporal serán anulados.

- **ENTREGA:**

- El proyecto es individual.
- Copias totales o parciales tendrán nota de 0 puntos y serán reportados a escuela.
- No se recibirán proyectos después de la fecha de entrega.
- No hay prórroga.

## ENTREGABLES

- Los siguientes archivos deben de ir dentro de una carpeta comprimida en formato **([BD1]Proyecto1\_#carnet.rar)**:
  - El modelo relacional propuesto en formato de imagen. Formato: **[BD1]ModeloPropuesto.png**.
  - Script (con extensión .sql) que contiene las instrucciones DDL para crear el modelo relacional. A este script debe de aplicarle todas las reglas de integridad vistas en clase y laboratorio. Formato: **[BD1]InstruccionesDDL.sql**.
  - Archivo de control utilizado para realizar la carga masiva desde el archivo CSV a la tabla temporal. Formato: **[BD1]ArchivoControl.ctl**.
  - Script (con extensión .sql) que contiene todas las consultas necesarias para llenar el modelo relacional propuesto a partir de la tabla temporal. Formato: **[BD1]CargaDeDatos.sql**.
  - Script (con extensión .sql) que contiene todas las consultas necesarias para realizar los reportes. Formato: **[BD1]Consultas.sql**.
  - Documentación (con extensión .pdf) que contiene todas las tablas, con su respectiva información de atributos, relaciones y explicación del modelo relacional. Formato: **[BD1]Documentacion.pdf**

## FECHA DE ENTREGA

- Martes 21 de Marzo de 2022, antes de las 11:59 p.m., a través de la plataforma de UEDI.