# Segundo parcial - Gestión de Datos

## 2do cuatrimestre de 2022

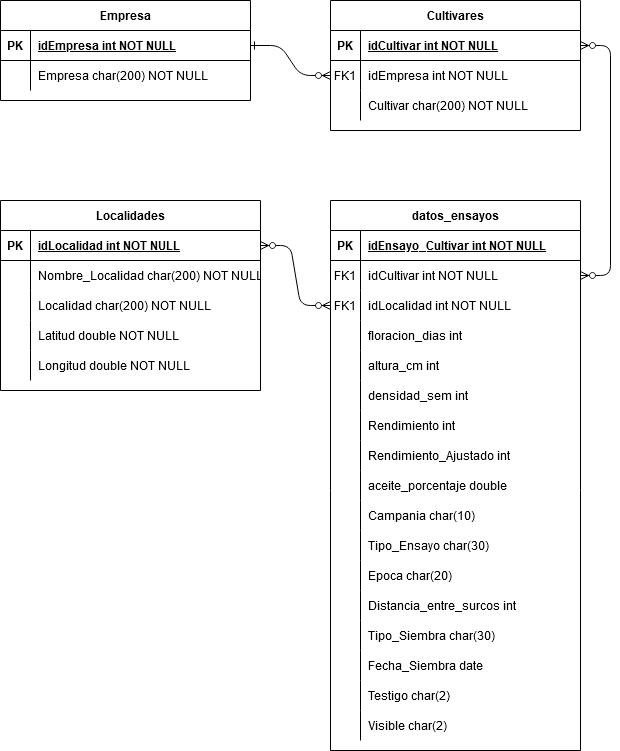
### Indicaciones generales

Deberán resolver las consignas propuestas de manera individual utilizando todas las herramientas y consultando los materiales vistos durante la materia. Al finalizar será necesario que suban la resolución del parcial como entrega al campus (solapa 2do parcial).

El parcial durará 2 (dos) horas y consta de dos partes. La primera se debe resolver con SQL y la segunda con R.

Para aprobar el parcial además de obtener una nota igual o superior a 4, deberán haber planteado ejercicios tanto de R como de SQL. Si resuelven una parte pero no la otra, el parcial se considera desaprobado.

### DER de la base de datos



### Primera Parte

1. En el campus buscar el archivo: **ensayos.db**, abrirlo en <https://sqliteonline.com/> para cargar la base de datos con cuatro tablas.
2. Para cada punto de esta primera parte guardar el SQL en **un solo archivo** como <apellido>\_*PrimeraParteParcial.sql*
3. Este es el DER de la base de datos, usarlo como apoyo para resolver los ejercicios.
4. Ejercicios:
   1. Escribir una consulta que muestre los ensayos de cultivares que tienen un rendimiento mayor a 5000. Mostrar el Id del cultivar, la floración, la altura, el rendimiento, el rendimiento ajustado y el porcentaje de aceite, ordenado por el Id del cultivar.
   2. Escribir una consulta que muestre el el nombre de la empresa y cuantos cultivares tiene. Ordenar la salida de mayor a menor con respecto de la cantidad de cultivares.
   3. Escribir una consulta que calcule los minimos y promedios de rendimiento, rendimiento\_ajustado, floracion, altura y porcentaje de aceite para cada Campaña.
   4. Escribir una consulta que muestre las distintas empresas que tienen ensayos en General Pico.
   5. Escribir una consulta que muestre una lista única de Localidades en las que se hacen ensayos donde el tipo de siembra es “Directa”.
   6. Listar cuantos ensayos testigos tiene cada localidad. Mostrar el nombre de la localidad y la cantidad. Ordenado de forma descendente por Cantidad.
   7. Generar un conjunto de datos que permita hacer un mapa de los lugares donde la empresa ‘ACA’ ha tenido ensayos. Para esto necesitamos la localidad y sus coordenadas geográficas (latitud y longitud). Incluir en ese conjunto de datos de campaña y la fecha de siembra.
   8. Obtener los nombres de las empresas que han hecho un solo ensayo.

### Segunda Parte

1. Para cada punto de esta segunda parte guardar el código de R en un solo archivo como <apellido>\_*SegundaParteParcial.R o .Rmd*
2. Entrar a RStudio y cargar los paquetes dplyr, dbplyr y RSQLite.
3. Cargar la base de datos ensayos.db y sus tablas que se encuentra en el campus usando el código para leer bases de datos utilizado en clase.
4. Escribir como código de R las consultas de los puntos: a, b, c y d de la primera parte.

| **SQL** | | **R** | |
| --- | --- | --- | --- |
| a | 0.5 | a | 0.75 |
| b | 0.5 | b | 0.75 |
| c | 0.5 | c | 0.5 |
| d | 1.0 | d | 1 |
| e | 0.75 | lectura de datos | 0.5 |
| f | 0.75 | **Total** | **3.5** |
| g | 1.0 |  |  |
| h | 1.5 |  |  |
| **Total** | **6.5** |  |  |