

Para el trabajo final del curso, utilizaremos un conjunto de datos llamado **agpop** disponible en el repositorio **SDaA** en CRAN. El recurso computacional a utilizar es de libre elección, para quienes no utilicen R, deberán exportar el conjunto de datos desde el repositorio CRAN para ser utilizado en el software escogido.

Para efectos de los cálculos deberán determinar una **semilla única** para su trabajo.

Descripción del conjunto de datos

La base de datos mencionada, corresponde a los resultados de un censo realizado por el gobierno de Estados Unidos mediante el departamento encargado de agricultura, el cual es realizado cada 5 años. En el estudio, se recolectan datos de todas las granjas (definidas como *cualquier lugar en el cual 1000 o más productos agrícolas son producidos y vendidos*) entre los 50 estados que componen la nación.

Los datos **agpop** contienen la información de los años 1982, 1987 y 1992 por condado y estado de la superficie total en acres dedicada a granjas, número de granjas, número de granjas menores a 9 acres y número de granjas mayores a 1000 acres, respectivamente según el orden de las columnas.

Utilizando el conjunto de datos descrito, deberá confeccionar un informe técnico que contenga lo siguiente:

1. Análisis exploratorio de datos del conjunto de datos.
2. Realización de un muestreo aleatorio simple de tamaño 400, del cual deberá confeccionar:
 - (a) Estimación del número total de acres dedicado a granjas en los Estados Unidos.
 - (b) Intervalo de confianza del 98 % para la cantidad anteriormente estimada.
 - (c) Análisis de la variación porcentual entre los años 1982, 1987 y 1992 del número total de acres dedicado a granjas en los Estados Unidos.
 - (d) Tamaño de muestra de muestra necesario con un 95 % de confianza tal que el error de estimación sea a lo más de 8.000 unidades.
3. Realización de un muestreo estratificado por estado de tamaño 400 (considerando cada estado como un estrato) usando una afijación proporcional y de Neyman respecto del número total de granjas. Utilizando los datos muestrales obtenido para cada afijación, realice:
 - (a) Estimación del número total de acres dedicado a granjas en los Estados Unidos.
 - (b) Intervalo de confianza del 95 % para la cantidad anteriormente estimada.
 - (c) Análisis de la variación porcentual entre los años 1982, 1987 y 1992 del número total de acres dedicado a granjas en los Estados Unidos.
 - (d) Análisis de la eficiencia relativa de la afijación Neyman con respecto de la afijación proporcional para cada año para estimar el número total de acres dedicado a granjas en los Estados Unidos.
4. Comparación de los resultados obtenido bajo un muestreo aleatorio simple y un muestreo estratificado.
5. Proponga una estimación indirecta para estimar el número total de granjas pequeñas (menores a 9 acres) para el año 1997. Justifique su propuesta en base a los años 1987 y 1992.
6. Conclusión.

El informe, presentación y códigos realizados deberán ser entregados a lo más el **miércoles 09/07** (antes de clase) mediante el módulo del curso que se pondrá a disposición. Las presentaciones se realizarán en las sesiones del **09/07** y **16/07** (en caso excepcional), deberán durar a lo más 1 hora.

Máximo de integrantes por trabajo: 2

La rúbrica de evaluación asociada los trabajos es la siguiente:

Criterios	Puntaje	Sobresaliente (5.5-7)	Suficiente (4.0 - 5.4)	Deficiente (< 4.0)
Estructura	1 - 7	El trabajo presenta claramente todos los contenidos solicitados.	El trabajo presenta claramente la mayoría de los contenidos solicitados.	El trabajo presenta menos de la mitad de los contenidos solicitados.
Metodología y resultados	1 - 7	La metodología usada es la correcta y los resultados son presentados claramente.	La metodología usada es la correcta y los resultados son presentados claramente con errores menores.	La metodología usada es incorrecta y los resultados son erróneos.
Conclusiones	1 - 7	Las conclusiones son presentadas claramente y demuestran una clara comprensión del material teórico.	Las conclusiones no son presentadas claramente y demuestran cierta comprensión del material teórico.	Las conclusiones son presentadas y no demuestran comprensión del material teórico.
Preguntas y presentación	1 - 7	Las preguntas son respondidas en su totalidad y la presentación es clara y fluida.	Las preguntas son respondidas de manera incompleta y la presentación es clara y mayoritariamente fluida	Las preguntas no son respondidas y la presentación no es clara ni fluida